Over the years, researchers have developed statistical methods to help them investigate and interpret issues of interest in many discipline areas. These methods range from descriptive to inferential to multivariate statistics. As the psychometrics measures in education become more complex, vigorous and robust methods were needed in order to represent research data efficiently. One such method is Structural Equation Modeling (SEM).

SEM is a statistical technique that allows the simultaneous analysis of a series of structural equations. It also allows a dependent variable in one equation to become an independent variable in another equation. It is a comprehensive statistical approach to testing hypotheses about relations among observed and latent variables. SEM is commonly known as causal modeling, or path analysis, which hypothesizes causal relationships among variables and tests the causal models with a linear equation system. As educational research questions become more complex, they need to be evaluated with more sophisticated tools. The pervasive use of SEM in the literature has shown that SEM has a potential to be of assistance to modern educational researchers.

This book brings together prominent educators and researchers from around the world to share their contemporary research on structural equation modeling in educational settings. The chapters provide information on recent trends and developments and effective applications of the different models to answer various educational research questions. This book is a critical and specialized source that describes recent advances in SEM in international academia.
Structural Equation Modeling
in Educational Research
CONTEMPORARY APPROACHES TO RESEARCH IN LEARNING INNOVATIONS

Series Editors:
Myint Swe Khine – Emirates College for Advanced Education, United Emirates
Lim Cher Ping – Edith Cowen University, Australia
Donald Cunningham – Indiana University, USA

International Advisory Board
Jerry Andriessen – University of Utrecht, the Netherlands, Kanji Akahori – Tokyo Institute of Technology, Japan, Tom Boyles – London Metropolitan University, United Kingdom, Thanasis Daradoumis – University of Catalonia, Spain, Arnold Depickere – Murdoch University, Australia, Roger Hartley – University of Leeds, United Kingdom, Victor Kaptelinin – Umeå University, Sweden, Paul Kirschner – Open University of the Netherlands, the Netherlands, Konrad Morgan – University of Bergen, Norway, Richard Oppermann – University of Koblenz-Landau, Germany, Jörg Zumbach – University of Salzburg, Austria

Rationale:

Learning today is no longer confined to schools and classrooms. Modern information and communication technologies make the learning possible anywhere, any time. The emerging and evolving technologies are creating a knowledge era, changing the educational landscape, and facilitating the learning innovations. In recent years educators find ways to cultivate curiosity, nurture creativity and engage the mind of the learners by using innovative approaches.

Contemporary Approaches to Research in Learning Innovations explores approaches to research in learning innovations from the learning sciences view. Learning sciences is an interdisciplinary field that draws on multiple theoretical perspectives and research with the goal of advancing knowledge about how people learn. The field includes cognitive science, educational psychology, anthropology, computer and information science and explore pedagogical, technological, sociological and psychological aspects of human learning. Research in this approaches examine the social, organizational and cultural dynamics of learning environments, construct scientific models of cognitive development, and conduct design-based experiments.

Contemporary Approaches to Research in Learning Innovations covers research in developed and developing countries and scalable projects which will benefit everyday learning and universal education. Recent research includes improving social presence and interaction in collaborative learning, using epistemic games to foster new learning, and pedagogy and praxis of ICT integration in school curricula.
# TABLE OF CONTENTS

Forward ix

## INTRODUCTION

Chapter 1  
**Modeling Educational Research: The Way Forward**  3  
Timothy Teo and Myint Swe Khine

## CONCEPTS AND METHODOLOGIES

Chapter 2  
**Assessing Spurious Interaction Effects**  13  
Andreas G. Klein, Karin Schermelleh-Engel, Helfried Moosbrugger and Augustin Kelava

Chapter 3  
**Latent Growth Model Analysis in Structural Equation Modeling: Concepts and Implementations**  29  
Xitao Fan and Timothy R. Konold

Chapter 4  
**Methodological Artifact or Substance? Examinations of Wording Effects Associated with Negatively Worded Items**  59  
Christine Distefano and Robert W. Motl

Chapter 5  
**Assessing Generalizability and Interpretability in Multilevel Data Structures with Unbalanced Multifaceted Measurements**  79  
Fadia Nasser Abu Alhija and Knut A. Hagtvet

Chapter 6  
**Testing Multiple Nonlinear Effects in Structural Equation Modeling: A Comparison of Alternative Estimation Approaches**  103  
Helfried Moosbrugger, Karin Schermelleh-Engel, Augustin Kelava and Andreas G. Klein
TABLE OF CONTENTS

Chapter 7
The Tenuousness of Invariance Tests Within Multisample Covariance and Mean Structure Models 137
Gregory R. Hancock, Laura M. Stapleton and Ilona Arnold-Berkovits

Chapter 8
Testing for Evidence of Construct Validity 175
Barbara M. Byrne

Chapter 9
Structural Equation Modeling in Educational Research 201
Kenneth J. Rowe

APPLICATIONS IN EDUCATIONAL RESEARCH

Chapter 10
Teachers’ Self-Reported Assessment Practices and Conceptions 243
Gavin T L Brown

Chapter 11
Comparing the Effect on the Development of Graduate Capabilities of the Differing Teaching and Learning Environments in Three Discipline Areas 267
David Kember and Doris Y.P. Leung

Chapter 12
Representing the Circles in Our Minds: Confirmatory Factor Analysis of Circumplex Structures and Profiles 287
Gabriel Nagy, Herbert W. Marsh, Oliver Lüdtke and Ulrich Trautwein

Chapter 13
Non-Linear Structural Equation Modeling 317
Fan Yang-Wallentin and Kajsa Yang-Hansen
TABLE OF CONTENTS

Chapter 14  
Measurement Invariance and Multi-Linear Group Analysis in SEM 329  
James B. Schreiber

Chapter 15  
Educational Applications of Latent Growth Mixture Models 345  
George A. Marcoulides and Ronald H. Heck

Chapter 16  
Parenting Quality and Child Externalizing Behavior: An Example of Longitudinal Panel Data Analysis 367  
Amanda Pearl, Brian F. French and Jean E. Dumas

List of Contributors 393
As budding researchers rise through the ranks of their statistical training, they gain familiarity and even comfort with the traditional hypothesis testing framework that pervades the practice of the general linear model. In transitioning to the world of structural equation modeling (SEM), however, particularly with the introduction of latent variables, they often find things becoming a bit unsettling, to say the least. Whereas their previous practice relied heavily on statistical tests and their resulting levels of significance, this new world starts with the notion that any model at hand is, in all likelihood, wrong. As if that weren’t enough, this modeling endeavor seems as much art as science, relying on a stream of researcher judgments throughout what is far more process than product. And in the end, no lone p-value sounds the trumpet of inference; rather, indices by the dozens present the researcher with that which is routinely far from crystal clear. Again, judgment is summoned, and with it the discomfort from its underemployment.

This process of SEM is, as described in the wonderful textbook by John Loehlin (2004), not unlike the crafting of an essay:

“So long as we want to try to describe complex real-life phenomena as they occur in their natural settings, it seems that our chief alternatives are the literary essay and the path model” (p. 232).

Indeed, SEM is a form of argument like the essay, resting on assumptions that may or may not be tenuous. The process is fraught with opportunity for our own biases and predispositions to color the avenues we choose, although we know intellectually that we must acknowledge and, if at all possible, resist their forces of attraction. And in the end, we arrange all of the evidence we have gathered, whose potential inconsistencies we must reconcile as best we can, in order to make our case.

Contributions to the SEM literature over the last three decades have really focused on enhancing our ability to make our case. Certainly much work has taken on the task of expanding this form of argument to be able to address new types of research questions and challenges. Other works have tackled the method’s underlying assumptions, either establishing when our inferences are robust to their violations or devising new processes that effectively circumvent such tenuous foundations. Considerable effort has also be directed, especially in the last decade or so, toward the clear communication of SEM and its best practices in order to ensure that our processes of argument are as immune as possible to our overly strong fore sights, and to our overly tempting hind sights.

The current volume represents a wonderful contribution to these continued efforts to strengthen the process of argument by SEM. Some chapters present new methods to address questions that are particularly challenging, such as those involving interactive or nonlinear relations. Some chapters ensure that we are aware of the cautions and limitations of inference, as in the case of multilevel data or multisample models. And the vast majority of chapters serve as exemplars for sound practice of SEM, utilizing fascinating questions across a wide variety of
disciplines as the context for modeling, so to speak, the practice of modeling. It has been a pleasure becoming acquainted with the current volume, and I am sure readers will share my enthusiasm.

Happy modeling!

Gregory R. Hancock
Professor and Chair
Department of Measurement, Statistics and Evaluation
University of Maryland

REFERENCE

INTRODUCTION
1. MODELING EDUCATIONAL RESEARCH: 
THE WAY FORWARD

INTRODUCTION

Structural Equation Modeling (SEM) as a statistical technique has increased in popularity since it was first conceived by Wright (1918), a biometrician who developed the path analysis method to analyse genetic theory in biology. SEM enjoyed a renaissance in the early 1970s, particularly in sociology and econometrics (Goldberger & Duncan, 1973), and later spread to other disciplines, such as psychology, political science, and education (Kenny, 1979; Werts & Linn, 1970). It was believed that the growth and popularity of SEM was attributed to a large part to the advancement of software development that have made SEM readily accessible to substantive researchers who have found this method to be well-suited to addressing a variety of research questions (MacCallum & Austin, 2000). Some examples of such software include LISREL (LInear Structural RELations), EQS, AMOS (Analysis of Moment Structures), and MPlus.

The combination of methodological advances and improved interfaces in various SEM software has resulted in the diverse usage of SEM. Hershberger (2003) examined major journals in psychology from 1994 to 2001 and found that over 60% of these journals contained articles using SEM, more than doubled the number of articles published from 1985 to 1994. Although SEM continues to undergo refinement and extension, it continues to be popular among applied researchers in other fields. Some examples include communication sciences (Holbert & Stephenson, 2002), operations management (Shah & Goldstein, 2006), and counseling psychology (Martens, 2005).

In recent years, the use of SEM has increased among educational researchers. To a large part, this could be attributed to the easy access to SEM literature and resources. A simple search on the internet would reveal numerous links to websites that contains teaching and learning materials, bibliographies, journal articles, blogs, book reviews, and free demo versions of SEM software, to name a few. Such preponderance of freely available resources serves not only to initiate new and potential users but also allow current and experience researchers to update themselves of the latest thinking and findings on SEM-related issues. As more educational researchers take to SEM as a tool to analyse complex relationships that exist, more dedicated resources are needed to meet their needs. This book aspires to be one such resource. It brings together the major thinkers and theorists in SEM who draws from their wealth of experience and knowledge to share ideas that are considered to be progressive in the application of SEM.
The first part of the book focuses on the concepts and methodologies of SEM. In chapter 2, Klein, Schermelleh-Engel, Moosbrugger, and Kelava highlighted the potential problems posed by spurious interaction effects in research. Although quadratic terms have been included in previous research to cope with spurious interaction effects, these were often recommended on an ad-hoc basis, hence a more systematic analysis of the problem was needed. In this chapter, the authors developed a methodological argument to show that the use of quadratic effects is sufficient to adjust for spurious interaction, under certain assumptions. They went on to show that, for latent variable models, using additional quadratic terms can provide a remedy against spurious interaction that is caused by non-normal predictor variables. This chapter concludes with an empirical example that illustrates the use of quadratic terms in combination with the LMS method for the analysis of latent interaction effects.

Chapter 3 by Fan and Konold provided a conceptual and procedural discussion on conducting latent growth modeling analysis within the framework of structural equation modeling (SEM). It is hoped that this chapter would provide the foundation for using SEM for latent growth modeling analysis, and serves as a springboard for further extensions of this quantitative approach for longitudinal study design. The authors discussed different latent growth models which include linear and non-linear growth models, a growth model with multiple growth series, a growth model with time-invariant covariates (predictors), and that with time-varying covariates (predictors). It was suggested that there are many other useful and creative variations and extensions of this quantitative approach, and collectively, they provide interesting analytic choices for longitudinal data analysis.

In chapter 4, DiStefano and Motl examined the wording effects that are associated with negatively worded items. Using the data from their work which investigates method effects due to negative wording using the RSES, the authors suggested that the method effect associated with negative wording showed the characteristics of a response style. In addition, the authors illustrated the utility of the CTCM model framework to separate the method effects from substantive content. This chapter has the potential to inform applied covariance modeling researchers on the existence of systematic effects attributable to item wording patterns that may be inherent in measures including both positively and negatively phrased items. By recognizing method effects, researchers could avoid misleading interpretations of scales in which both positively and negatively worded items are utilized. The authors concluded by proposing more work to be done in this area in order to gain a greater understanding of the meaning and causes of method effects.

Chapter 5 assessed the generalizability and interpretability in multilevel data structures with unbalanced multifaceted measurements. Although Cronbach’s alpha is a common way to estimate reliability in one-facet one-level measurement designs, Nasser and Hagtvet argued that confirmatory factor analysis is a proper choice for testing unidimensionality. However, estimation of reliability in a multilevel data structure with one- or multifacet measurement designs is less known, with ways of assessing the structure of universe score variance in multilevel data designs where
multifacet test designs to be even less familiar. This chapter attempts to illustrate a methodological approach to generalizability assessment as well as to interpretability in such a data structure. This two-fold objective calls for a joint application of generalizability analysis and confirmatory factor analysis.

In chapter 6, Moosbrugger and colleagues explained the nonlinear effects that exist when the size or direction of the effect of a predictor variable on a criterion variable varies with the level of a moderating variable (interaction effect) or with the level of the predictor itself (quadratic effect). While latent interaction effects have received considerable attention in the methodological literature, relatively few studies have been conducted that analyze both interaction and quadratic effects simultaneously in a structural equation model. This chapter aims to compare the performance of four alternative estimation approaches for the simultaneous analysis of multiple latent nonlinear effects. The results of the simulation studies reveal that LMS and QML outperform both LISREL approaches and that the constrained approach performs better than the unconstrained approach when multicollinearity is in the data.

The tenuousness of invariance tests within multisample covariance and mean structure models was explicated by Hancock and colleagues in chapter 7. In light of the growing use of multigroup latent variable models, the authors addressed challenges faced when conducting tests of invariance within both covariance and mean structures. Specifically, they demonstrated how all tests of invariance are conditional upon the fundamental and untestable assumptions embedded in the constraints required for model identification by considering the choice of each factor’s scale indicator for the covariance structure and the single required intercept constraint for the mean structure. Hancock and colleagues concluded that no sequence of tests can ensure that one’s assessment of invariance matches the true population state of invariance, and when a mismatch occurs then tests of structural and mean parameters can become greatly compromised.

In chapter 8, Byrne shared on the testing for evidence of construct validity by using two approaches to the Multitrait-Multimethod model. Using the MTMM matrix in determining the extent to which a measure provided evidence of construct validity, Campbell and Fiske proposed qualitative decision rules based on the following three concepts: (a) convergent validity, the extent to which different assessment methods concur in their measurement of the same trait; (b) discriminant validity, the extent to which independent assessment methods diverge in their measurement of different traits; and (c) method effects, an extension of the discriminant validity issue, represent bias that can derive from use of the same method in the assessment of different traits; correlations among these traits are typically higher than those measured by different methods. Byrne concluded by suggesting that, in addition to technical considerations such as convergence to proper solutions and goodness-of-fit, researchers should place a heavy emphasis on substantive interpretations within the related theoretical framework.

Part 1 of this book ends with a discussion on the use of SEM in educational research, with an emphasis on taking account of the measurement, distributional and structural properties of the data by fitting multilevel structural equation
models. In chapter 9, Rowe advocated the need to consider the inherent clustered or hierarchical structure of the data typically obtained in educational research (e.g., students within class/teacher groups, within schools, etc.). Using an annotated example, this chapter demonstrates the utility of fitting multilevel structural equation models to a hierarchically structured data set consisting of 4527 students in 339 class/teacher groups, drawn from a two-stage, cluster-designed PPS sample of 52 elementary schools to estimate the simultaneous interdependent effects operating among students’ Literacy achievements, inattentive behaviors and attitudes to school – both at the student-level and at the class/teacher-level.

APPLICATIONS IN EDUCATIONAL RESEARCH

The chapters in Part II delve with applications of SEM in educational research. SEM as one of the statistical procedures is relatively new approach. However over the last two decades computational and statistical power has grown tremendously and a number of software programs have been developed and proved that this approach is a viable option in dealing with latent variables.

Brown’s study in Chapter 10 involves teachers’ self-reported assessment practices and conceptions. He explored New Zealand primary school teachers’ conceptions of assessment and frequency of assessment practices by using self-administered questionnaire. In his literature review he found that teachers reported multiple and sometime conflicting conceptions of assessment. Some teachers indicated that assessment can improve teaching and learning, but others see that assessment is irrelevant. In some studies teachers have argued that assessment interferes with their work and teachers disagree using assessment to make student accountable. With this backdrop he attempted to connect conceptions to practice about assessment in his context. The analysis was done using structural equation modeling and the results found two acceptably fitting models for teachers’ self-reported conceptions of assessment and frequency of assessment practices. Exploratory factor analysis (EFA) was also carried out on the assessment practices questionnaire to test the model of two cognitive processing factors and the four assessment practice factors. EFA indicates the existence of four factors. In the structural model, he found that all assessment practices correlated positively with all assessment purposes. He concluded that the teachers conceived of assessment for the purpose of improvement and student accountability, but not to use assessment for surface learning. He noted the power of structural equation modeling to detect structures and relations within the thinking of teachers about educational assessment.

David Kember and Doris Leung presented their findings on the effect on the development of graduate capabilities of the differing teaching and learning environments found in three discipline areas namely; business administration, engineering and science (hard sciences) and the remainder (grouped together as arts, education, health sciences and social sciences). Students completed the Student Engagement Questionnaire which comprises 15 scales; out of which six scales measure students’ perceptions of the development of six generic capabilities
and nine scales measure the teaching and learning environment. The six capability scales comprised of (i) critical thinking, (ii) self-managed learning, (iii) adaptability, (iv) problem solving, (v) communication skills and (vi) interpersonal skills and group work. The nine elements in the teaching and learning environment consisted of (i) active learning, (ii) teaching for understanding, (iii) feedback to assist teaching, (iv) assessment, (v) teacher/student interaction, (vi) assistance from teaching staff, (vii) student/student interaction, (viii) cooperative learning and (ix) coherence of curriculum. The study built upon previous work which used SEM to investigate the impact of the teaching and learning environment upon the development of generic capabilities. The model hypothesized the nine scales of the teaching and learning environment influencing the six capabilities. The total of 3305 usable data sets were obtained. EQS software was used for multiple-group structural equation modeling. The results indicated that students in the remainder group perceived a more positive teaching and learning environment than students in the hard sciences and the business group. Latent mean difference suggest statistical differences in the two capabilities latent constructs across three groups. The authors concluded that the use of multiple-group SEM established the fact that the same model applied to the three disciplinary groups with differences in the magnitude of structural paths and latent means. This was interpreted as indicating differing approaches to teaching between disciplines.

In Chapter 12, Nagy, Marsh, Ludtke and Trautwein present a confirmatory factor model for the circumplex and address factor analytic methods for simultaneously representing correlation patterns and individual profiles. The authors noted that a circumplex can be characterized as a spatial representation of individual differences. They described examples for applications of circumplex models in determining the interpersonal behavior, personality traits, emotions, values and so on. The paper outlined two aspects of the circumplex, i.e. variable-oriented perspective of correlation patterns, and the person-oriented perspective of profiles. The authors showed how factorial representations relate these two aspects. The chapter proved the use of circumplex models and use of sophisticated and flexible methods to examine circumplex hypotheses.

Nonlinear equation modeling and analysis of reading literacy data in Sweden was presented by Fan Yang-Wallentin and Kajsa Yang-Hansen in Chapter 13. The chapter began with the issues related to interactive or nonlinear effects on outcome of latent variable and properly performing structural equation modeling. The authors explained the Latent Variable Scores (LVS) as one of the methods available for estimating latent variables. The data used to test the hypothetical model is from the International Association for the Evaluation of Educational Achievement (IEA) to test the reading ability in over 30 educational systems. The sample is taken from the reading achievement test of Swedish 4th grade students in Reading Literacy Study 2001. It involved 5361 students with the age ranging from 9 to 10 years old.

Studies in the past documented that reading achievement is determined not only by the general reading ability and specified reading skills such as document reading skills and besides decoding/reading speed also attributes to the variation in reading achievement. They hypothesized that a nonlinear development of reading
comprehension may also underlie in the data. In this chapter the authors presented the use of LVS procedure by Anderson & Robin (1996) indicating that the latent variable scores have the same covariance matrix as the latent variable themselves. When testing the measurement model for Reading Comprehension, Linguistics Cognition and Reading Speed, it was found that the model did not fit data exactly, but close enough to assume that the measurement model is valid. The results showed that there is an interaction effect of Linguistic Cognition and Reading Speed on Reading Comprehension. The authors concluded that the reading comprehension examples showed that LVS is a simple and easy to operate method.

Chapter 14 deals with the measurement of invariance through multi-group confirmatory factor analysis (MGCFA) along with secondary analysis using multiple indicator multiple causes (MIMIC) model. In this chapter James Schreiber reported an examination of a brief student course evaluation form used in Duquesne University in the USA. By using measurement invariance and multi-group analysis in SEM, he presented the research findings from using brief eight-item student evaluation instrument. The data were collected from 1138 students to assess the Teaching Quality and Positive Learning Environment. Schreiber indicated that “teacher quality” is a latent construct that is composed of many variables. SEM is more appropriate dealing with such latent variables. He explained that “SEM and ANOVA/MANOVA treat measurement errors differently. In SEM the latent constructs are assumed error free because error term is associated with the observed variable”.

The original instrument used to measure teacher quality contained 55 items with 10 latent constructs. The duration to complete the questionnaire posted problem to the students and the questionnaire was revised and shortened to 8 items. This brief instrument contained 5 items on Teaching Quality and 3 items on Positive Learning Environment. The information about student’s perception of high or low class performance expectations before final examination were also collected. The data was analyzed using AMOS 4.0 Graphics. The results found that there is a difference between low and high performance. The group who perceived themselves as high performers had a more positive impression of the learning environment. The author noted that in teaching and learning domain the use of simple exploratory factor analysis and internal consistency values on the questionnaires are no longer adequate. It is important to conduct multi-group analysis to ensure that the instrument functions across groups. This chapter confirmed that SEM is a flexible family of analysis techniques and one of them is the analysis of data from multi-groups.

Marcoulides and Heck presented how latent growth mixture models (LGMM) can be used in education. Chapter 15 introduces the conceptual and methodological details related to the use of latent growth mixture modeling and describes a model of academic achievement level and growth in mathematics. In their chapter Marcoulides and Heck indicated the existence of new types of mixture models such as latent growth mixture models, multilevel growth mixture models, latent transition analysis models, Markov chain models and latent variable hybrid models. These models expand the range of developmental processes that can be examined in the variety of settings.
To illustrate the application of LGMM, they have examined a single-level growth mixture model for mathematics achievement of 6623 students using three within-school covariates. The results suggest that students with higher Grade Point Averages (GPA) and higher Socioeconomic Status (SES) have significantly higher initial status than their peers with lower GPAs and SES. They also found that students with higher GPAs and SES make significantly greater growth over time than peers with lower GPAs and SESs. The result of the model was tested and it suggests that class membership was partially explained by student SES and GPA. They also reported that between schools, social composition was significantly related to initial status achievement level, but not related to growth rate. The author concluded that complexity of LGMM varies, but these kinds of models offer possibilities for applied researchers in attempting to answer substantive theories.

In Chapter 16 Pearle, French and Dumas presented an example of longitudinal analysis based on the data on parenting quality and child externalizing behavior. The study made the reference to the questions of whether the child development including educational settings shapes the parents their children or children their parents. The authors argue that a unidirectional model cannot explain the subsequent outcome in either child adjustment or parenting quality. Multiple-wave longitudinal data that represents the collection of behaviors over time on the same persons and same variables is analyzed using SEM. The data set involved 249 children enrolled in Kindergarten. They have noted that “this framework not only allows for the examination of changes in mean levels of outcome variables over time, but also important other model parameters that may influence the mean change”.

Child Behavior Checklist (CBCL) was used to measure the child adjustment and Disruptive Behavior Disorder Rating Scale (DBD) was used to measure the child externalizing behavior. Primary caregivers responded to the questionnaires. The authors assessed the measurement models and cross-legged models in a latent variable framework (SEM). The results found that each latent construct, parenting quality and child externalizing behavior was fit individually. They reported that measurement models for both latent constructs were also fit over time. The authors concluded that wide variety of software packages are available, but modeling strategy is the same. They suggested using similar methodology as explained in the chapter when the research is concerned with the assessment of change in constructs over time or across conditions.

The chapters in Part II demonstrated that the use of Structural Equation Modeling in educational research is a powerful multivariate analysis and flexible. It also served as an extension of the general linear model and specialized software that can be used for special cases. It is hope that the case studies and examples described in these chapters will be useful resource for further educational research using SEM.

CONCLUSION

Since Wright’s (1918) introduction of path analysis method, the use SEM as statistical approach has gained momentum and progressed steadily over the past years. The measures used in today’s educational research have become more
complex hence vigorous and robust methods are needed in order to analyse research data efficiently. Many educational researchers have started using Structural Equation Modeling (SEM) and the outcomes of their analyses have helped to determine the effectiveness of learning innovations in various educational contexts. To consolidate the theoretical perspectives and advances in applications, researchers have shared the use of structural equation modeling and their recent findings in this book. It is hoped that the chapters in this book would provide information on trends and developments and effective applications of the different models to address various educational research questions.

REFERENCES


Timothy Teo
*Nanyang Technological University*
*Singapore*

Myint Swe Khine
*Emirates College for Advanced Education*
*United Arab Emirates*
CONCEPTS AND METHODOLOGIES
2. ASSESSING SPURIOUS INTERACTION EFFECTS

ABSTRACT
Spurious interaction effects have been an issue of concern in regression analysis, and the inclusion of quadratic terms has been advised in previous research to cope with that problem. However, some of these remedies were recommended on an ad-hoc basis, and a more systematic analysis of the problem was lacking. In this paper, a theoretical argument is developed which shows that under certain assumptions the use of quadratic effects is indeed sufficient to adjust for spurious interaction. Spurious interaction is further studied for latent variable models, and it is shown that the use of additional quadratic terms can provide a remedy against spurious interaction that is caused by nonnormal predictor variables. An empirical example is presented that illustrates the use of quadratic terms in combination with the LMS method for the analysis of latent interaction effects.

INTRODUCTION
Over the last two decades, structural equation modeling (SEM) has become a common statistical tool for modeling relationships between variables which cannot be observed directly, but only with measurement error. The relationships between these unobservable, latent variables are formulated in structural equations, and they are measured with errors by indicator variables in a measurement model. By the development of software packages for covariance structure analysis such as Amos (Arbuckle, 1997), EQS (Bentler, 1995; Bentler & Wu, 1993), LISREL (Jöreskog & Sörbom, 1993, 1996), and Mplus (Muthén & Muthén, 1998-2007), SEM has become available to a large community of researchers.

In a social science research context, a linear model structure sometimes provides only a questionable representation of reality. This is particularly the case when in a cross-sectional investigation the relationship between a predictor and the criterion variable itself depends on the outcome of a third variable. While ordinary SEM incorporates linear relationships among latent variables, models with nonlinear structural equations have attracted increasing attention over the last decade. Several researchers have called for estimation methods for nonlinear latent variable models, and numerous substantive theories in education and psychology call for analysis of nonlinear models (Ajzen, 1987; Ajzen & Fishbein, 1980; Ajzen & Madden, 1986; Cronbach, 1975; Cronbach & Snow, 1977; Fishbein & Ajzen, 1975; Karasek, 1979; Lusch & Brown, 1996; Snyder & Tanke, 1976). Also, a need...
for nonlinear extensions of ordinary SEM has been expressed from a methodological perspective (Aiken & West, 1991; Busemeyer & Jones, 1983; Cohen & Cohen, 1975; Jaccard, Turisi & Wan, 1990), and different estimation approaches have been proposed: Approaches that involve the forming of products of indicators (Hayduk, 1987; Bollen, 1995, 1996; Jaccard & Wan, 1995; Joreskog & Yang, 1996, 1997; Marsh, Wen & Hau, 2004), a moment-based 2-step estimation technique (Wall & Amemiya, 2003), a maximum likelihood estimation (Klein & Moosbrugger, 2000), and an approximate ML estimation (Klein & Muthén, 2007).

With the LMS (latent moderated structural equations) method, Klein and Moosbrugger (2000) introduced a maximum likelihood estimation technique for latent interaction models with multiple latent product and quadratic terms. LMS does not require the researcher to form product indicators for the latent product terms, and it takes the mathematical structure of the nonnormal distributions that are implied by the product terms into account. In this approach, the latent independent variables and the error variables are assumed to be normally distributed. In contrast to non-ML estimation methods, LMS also allows for likelihood ratio model difference tests which can test for the significance of one or several nonlinear effects simultaneously. LMS has been implemented in the Mplus structural equation modeling software (Muthén & Muthén, 1998-2007).

This paper addresses the problem of spurious interaction effects, that is, interaction effects that are falsely identified by statistical analysis although the true structure of the relationship does not involve interaction. The problem has been identified in the context of nonlinear regression analysis and studied by several researchers (Lubinski and Humphreys, 1990; Shepperd, 1991; Cortina, 1993; Ganzach, 1997). Since Saunders’ (1956) article on interaction, interaction effects in linear regression have mostly been modeled by adding a product term \( XZ \) of two predictor variables \( X \) and \( Z \) to a linear model equation: Conditional on a value \( z \) of one predictor (e.g., \( Z = z \)), the relationship between the other predictor (e.g., \( X \)) and a criterion variable \( Y \) is linear, and the slope of \( X \) on \( Y \) depends linearly on the value of the first predictor. In this context, the predictor \( Z \) has also been called a moderator variable that moderates the slope of the relationship of the second predictor \( X \), although the regression model itself is symmetric in \( X \) and \( Z \), and does not indicate what variable is counted as the moderator. In several simulation and case studies, it was found that a basic regression model with regressors \( X \), \( Z \), and \( XZ \) involves serious problems related to the interpretation and estimation of the regression coefficient for the product term. Lubinski and Humphreys (1990) found that when \( X \) and \( Z \) are correlated, quadratic terms \( X^2 \) and \( Z^2 \), when entered into the interaction model, can considerably change or even neutralize the unique variance explained by the product term, because of multicollinearity between the product term \( XZ \) and the quadratic terms. Depending on the signs of the quadratic and product terms in a true model, regression coefficients for these nonlinear terms could be over- or underestimated when some of the terms are omitted in the analysis. Ganzach (1997) showed that the omission of quadratic terms can both affect the Type I and the Type II error for the test of the significance of the product term. Following Cortina (1993), he advocates a conservative standpoint when testing for interaction. He recommends to include quadratic terms when testing for
interaction, and to employ a sequential testing procedure to avoid spurious interaction effects: First, a regression model with linear and quadratic terms is fit (i.e., with \( X, Z, X^2, \) and \( Z^2 \) as regressors), and then, in a second step, the product term \( XZ \) is added to test for interaction.

For this paper, we also adopt a conservative standpoint towards interaction effect analysis, that is based on the idea that an interaction model should only be adopted when it considerably improves modeling above and beyond a simpler, non-interactive modeling structure. Behind the more technical matters discussed by the abovementioned authors, as to what degree and under what exact conditions models with partially omitted quadratic or product terms can lead to biased or misleading estimates, there lies the more conceptual question about the adequacy of nonlinear terms to represent what is generally understood to count as an interaction effect. In particular, this leads to the following questions: First, given the general idea of an interaction effect, what recommends the selection of the product term to model interaction? And, second, once the product term is used to model interaction, is the inclusion of quadratic terms in the model sufficient to reduce the risk of a spurious interaction effect, or could it be that other regressors also need to be included to avoid that risk? These questions will be addressed in Section 1 (Interaction effects, additive models, and the role of the product term) and Section 2 (Use of quadratic baseline model to avoid spurious interaction in regression). In Section 1, we will start with a general notion of an interaction effect, where we approximate the regressive relationship by a polynomial and explain why the product term plays a unique role for modeling interaction. In Section 2, we deliver a theoretical argument for the sufficiency of the quadratic baseline model which shows that under linearity assumptions about the relationship between the predictors \( X \) and \( Z \), no further regressors beyond the linear and quadratic terms are needed for a non-interactive baseline model. In Section 3 (Use of quadratic baseline model to avoid spurious latent interaction effect), we look at latent variable interaction and conduct a simulation study that demonstrates how the inclusion of quadratic terms in a latent variable model can considerably reduce the risk of spurious interaction, especially when the distributional assumptions of the latent variable model are violated. Finally, in Section 4 (Empirical example), we analyze behavioral data to illustrate how a quadratic baseline model might be used to test for a latent interaction effect and help reduce the risk of an interaction artifact.

INTERACTION EFFECTS, ADDITIVE MODELS, AND THE ROLE OF THE PRODUCT TERM

An interaction effect of two predictor variables \( X \) and \( Z \) on a criterion variable \( Y \) is characterized by the fact that the relationship between one predictor (e.g., \( X \)) and the criterion \( Y \) changes across the levels of the other predictor (e.g., \( Z \)). This general notion of an interaction effect can be translated into a more precise general definition when one regards the conditional expectation function \( E[Y| X=x, Z=z] \). Conditional on certain values \( x \) and \( z \) for the predictors \( X \) and \( Z \), \( E[Y| X=x, Z=z] \) is the expected mean of the criterion variable \( Y \). \( E[Y| X=x, Z=z] \) is a bivariate
function of x and z, and it is the generally unknown regression function that describes the true mean relationship between Y, X, and Z in the population.

Stolzenberg (1980) distinguished between additive and non-additive regression functions, and this distinction leads us directly to a general definition of an interaction effect. The regression function $E[Y|X=x, Z=z]$ is additive, if it can be decomposed in the following way:

$$E[Y|X=x, Z=z] = u(x) + v(z), \quad (1)$$

with certain functions $u$ and $v$. Such an additive mean relationship is identified as a relationship without interaction: Across different levels $Z = z_1$, $Z = z_2$ of the variable $Z$, the mean relationship between $Y$ and $X$ is described by two “parallel” regression functions $u(x)+v(z_1)$, $u(x)+v(z_2)$, which differ only by a certain constant. Conversely, if for any two levels $Z = z_1$, $Z = z_2$ of $Z$ the two regression functions $E[Y|X=x, Z=z_1]$ and $E[Y|X=x, Z=z_2]$ only differ by a certain constant, $E[Y|X=x, Z=z]$ must be additive. This observation now allows us to characterize an interaction effect in a very general way: By definition, predictors $X$ and $Z$ interact in their relationship with a criterion variable $Y$ when the regression function $E[Y|X=x, Z=z]$ is not additive.

This definition provides a general notion of an interaction effect that is not bound to a particular parametric model. However, since the shape of the regression function $E[Y|X=x, Z=z]$ is generally unknown, for an ordinary regression analysis of specific data a parameterized structure (e.g., a linear model) needs to be assumed for the regression equation that models the function $E[Y|X=x, Z=z]$. This leads to the question what regressors and mathematical terms should be included in such a regression model, which is intended to test for an interaction effect.

Since Saunders’ (1956) introduction of interaction effects in regression analysis, these effects have usually been analyzed by including a product term $XZ$ of the interacting predictors $X$ and $Z$ in the model equation:

$$y_i = \alpha + \gamma_1 x_i + \gamma_2 z_i + \gamma_3 x_i z_i + \epsilon_i = (\alpha + \gamma_2 z_i) + (\gamma_1 + \gamma_3 z_i) x_i + \epsilon_i. \quad (2)$$

The equation describes a symmetric relationship where, across the levels of one predictor (i.e., across levels $Z=z$), the regressive relationship between $Y$ and $X$ is linear and its slope changes as a linear function of that first predictor $Z$. The model in Equation (2) is not additive, and it models an interaction effect when the parameter $\gamma_3$ is nonzero. But besides the feasibility and simple structure of this regression model, the question remains under what conditions the variance increment explained by the product term $XZ$ can be attributed to interaction, and whether the linear baseline model that is part of Equation (2) (i.e., the model resulting when $\gamma_3$ is zero) is an appropriate implementation of an additive model that serves as a model to represent the null hypothesis of no interaction.

Apart from an ad-hoc inclusion of the product term into the model, its critical role for modeling interaction can be seen when one regards a series expansion of the regression function. For this purpose, we assume that, without loss of generality,
the predictors X and Z are variables with zero means. Under the assumption that
the regression function \( f(x,y) = E[Y| X=x, Z=z] \) is a continuous, differentiable
function of \( x \) and \( z \), a second-degree Taylor approximation of \( f \) at the means of the
predictors (at \( x=0, z=0 \)) yields:

\[
E[Y| X=x, Z=z] \approx f(0,0) + \left[ \frac{\partial f}{\partial x} (0,0) \right] x + \left[ \frac{\partial f}{\partial z} (0,0) \right] z \\
+ \left[ \frac{\partial^2 f}{\partial x \partial x} (0,0) \right] x^2 + \left[ \frac{\partial^2 f}{\partial z \partial z} (0,0) \right] z^2 + \left[ \frac{\partial^2 f}{\partial x \partial z} (0,0) \right] xz.
\]  

(3)

A regression model that adopts this second-degree approximation then includes
both linear and quadratic terms, as well as the product term \( XZ \):

\[
Y = \alpha + \gamma_1X + \gamma_2Z + \gamma_3X^2 + \gamma_4Z^2 + \gamma_5XZ + \varepsilon.
\]  

(4)

In the approximation in Equation (3), the expressions in the brackets are
coefficients that depend on the unknown regression function \( E[Y| X=x, Z=z] \).
The approximation contains an intercept \( f(0,0) \), two linear terms, two quadratic
terms, and a product term \( xz \). The product term is the lowest-degree term in the
series expansion that approximates any non-additive structure that is part of
\( E[Y| X=x, Z=z] \).

The point \( (0,0) \) represents the mean vector of the two predictors. If, in the
neighborhood of the point \( (0,0) \), the slope of the regression function of \( Y \) on one
predictor (i.e., \( X \)) changes across the levels of the other predictor (i.e., across levels
\( Z=z \)), then the coefficient in front of the product term \( xz \) in Equation (3) is nonzero,
and \( E[Y| X=x, Z=z] \) is non-additive. Conversely, if \( E[Y| X=x, Z=z] \) is additive in
the neighborhood of the point \( (0,0) \), then the coefficient in front of the product term
\( xz \) is zero, and Equation (3) reduces to a purely additive function.

Altogether, this shows that the product term \( xz \) plays a key role for modeling
interaction in the regression function \( E[Y| X=x, Z=z] \). The product term is the
lowest-degree approximating term in the expansion that can model a non-additive
structure of the regression \( E[Y| X=x, Z=z] \), and if there is a change of slopes in the
neighborhood around the mean vector of the predictor variables, the coefficient in
front of the product term \( xz \) in Equation (3) will be nonzero.

**USE OF QUADRATIC BASELINE MODEL TO AVOID SPURIOUS INTERACTION
IN REGRESSION**

In the last section, we saw that the inclusion of a product term into a linear
regression equation is essential for modeling a non-additive structure of the
regression function \( E[Y| X=x, Z=z] \), and hence essential for modeling an
interaction effect. Linear regression analysis maximizes the explained variance for
the criterion variable, for a linear function of given regressors. For regression
analysis of interaction effects, this poses the following practical question: When
testing interaction, what regressors need to be included in an additive, non-
interactive baseline model, so that the additional variance explained by the product
term can be safely attributed to an interaction effect? Lubinski and Humphreys
(1990) as well as Ganzach (1997) have done significant research in this direction.
They conducted data analyses and simulation studies of regression models with a
product term XZ in addition to linear and quadratic terms in the model equation.
They found that the quadratic terms and the product term can be correlated, and
that the inclusion of the quadratic terms in the baseline model can considerably
reduce the variance increment of the product term when the latter is added to the
baseline model. If the true regression E[Y| X=x, Z=z] is an additive function with
linear and quadratic regressors, a regression analysis that only uses linear regressors
and a product term can lead to a spurious interaction effect, because part of the
variance of the quadratic regressors may now be explained by the product term. In
this context, Shepperd (1991) points out that this problem particularly occurs when
the predictors X and Z are correlated: The higher X and Z are correlated, the
higher usually is the correlation among the product term XZ and the squares X²
and Z² (see also Moosbrugger, Schermelleh-Engel, Kelava, & Klein, 2008). The
recommendation that follows from these studies is that, for a conservative
approach to interaction effect analysis, both linear and quadratic terms should be
included in the additive baseline model. Unless either quadratic effects can be
excluded on theoretical grounds or the predictors X and Z are uncorrelated,
interpreting a significant product term as evidence for an interaction effect may be
misleading and incorrect.

In ordinary regression analysis, no distributional assumptions about the regressor
variables are made. For the analysis of latent interaction effects, however, the
inclusion of quadratic terms might be of further relevance, because for latent
variable models additional distributional assumptions are often made which, when
violated, may lead to spurious interaction effects for other reasons than in
regression analysis. This aspect will be investigated in Section 3.

After following the ad-hoc recommendation to include quadratic terms into the
baseline model, the question arises whether the inclusion of the quadratic terms is
all sufficient. Could it be that additional regressors that are either functions of X or
Y, such as powers of X or Z, should be included in the regression model to avoid a
spurious interaction effect? Could it be that the inclusion of such additional
regressors changes the regression coefficient for the product term? To answer this
question on the population level, again, a theoretical argument can be provided. We
show that under a linearity assumption about the mean relationships between X and
Z, the additional regressors, when added to the regression model in Equation (4),
would not alter the regression coefficient γ_5 for the product term XZ. To see this,
we assume that both X and Z have zero means, and that the mean relationships
E[Z| X=x] and E[X| Z=z] are linear. Under the assumption that E[Z| X=x] is linear,
we then have:
ASSESSING SPURIOUS INTERACTION EFFECTS

\[ E[Z \mid X=x] = b \times x \text{ and } Z = bX + \varepsilon, \quad (5) \]

where \( \varepsilon = Z - bX \) is the residual with \( E[\varepsilon] = 0 \). \( \varepsilon \) is further assumed to be statistically independent from \( X \). Then we obtain for the product term \( XZ \):

\[
\text{Cov}(g(X), XZ) = \text{Cov}(g(X), bX^2 + \varepsilon X) \\
= \text{Cov}(g(X), bX^2) + \text{Cov}(g(X), \varepsilon X) \\
= \text{Cov}(g(X), bX^2) + E[g(X) \varepsilon X] \\
= \text{Cov}(g(X), bX^2), \quad (6)
\]

where \( g(X) \) is any arithmetic function of \( X \). Thus, \( g(X) \) and \( XZ \) share no common variance except for a variance component that is already fully accounted for by \( X^2 \). Similarly, under the assumption that \( E[X \mid Z=z] \) is linear and the residual of this is independent from \( Z \), it follows that \( g(Z) \) and \( XZ \) share no common variance except for a variance component that is already fully accounted for by \( Z^2 \). Altogether, we can conclude that, under the linearity assumptions made, adding a regressor \( g(X) \) or \( g(Z) \) to Equation (4) would not change the regression coefficient for the product term \( XZ \). Under the assumptions made, once the linear and quadratic terms are used in the additive baseline model, no further regressors seem to be necessary to extend the additive baseline model. The variance increment of the product then explains a variance component of \( Y \) that cannot be accounted for additively by \( X \) and \( Z \). The variance increment of the product term can then be genuinely attributed to an interaction effect. This theoretical argument goes beyond the ad-hoc recommendation that was suggested in the previous literature. It supports the idea to use an additive baseline model that has linear and quadratic regressors in it, but no other types of regressors.

USE OF QUADRATIC BASELINE MODEL TO AVOID SPURIOUS LATENT INTERACTION EFFECT

In the previous section, we provided a recommendation for the analysis of interaction effects in the framework of regression analysis. Based on theoretical considerations, our findings support the idea of using a product term to model an interaction effect, and they support the idea to use an additive baseline model with linear and quadratic terms. In a linear regression analysis with ordinary least squares estimation, only the first and second moments of the variables (i.e., means, variances, and covariances) are used for the estimation of the regression coefficients and standard errors. Estimates for the regression coefficients, for instance, are not affected by skewness of the regressors, as long as the linearity of the relationship between the dependent variable and the regressors holds. In particular, there are no distributional assumptions made for the regressor variables. But the situation is very different for parameter estimation in latent variable models, where, for example, ML estimation is applied. Then, usually additional distributional assumptions such as the assumption of multivariate normality of the indicator variables are made, and
a violation of these assumptions might be critical for the estimates and the standard errors of such models.

In this section, we investigate the problem of spurious interaction effects for latent variable models. Specifically, we will look at the LMS method (Klein & Moosbrugger, 2000) for the analysis of latent interaction effects. The LMS method has been implemented in the Mplus (Muthén & Muthén, 1998-2007) structural equation modeling software since 2004. It provides an ML estimation of the model parameters under the assumption that the exogenous variables in the model (latent predictor variables, residual variables, and measurement errors) are normally distributed. LMS has been shown to provide very efficient estimators when the distributional assumptions are met (Schermelleh-Engel, Klein & Moosbrugger, 1998; Klein & Moosbrugger, 2000; Dimitruk, Schermelleh-Engel, Kelava, & Moosbrugger, 2007; Kelava, Moosbrugger, Dimitruk, & Schermelleh-Engel, 2008). On the other hand, in simulation studies with skewed latent predictor variables, there was indication for some possible problems related to spurious interaction effects and inflated Type I error when testing for the significance of a latent product term (Klein & Moosbrugger, 2000; Klein & Muthén, 2007).

However, in the studies where these specific problems were first found, no quadratic terms were used, and a latent product term was tested against a purely linear baseline model. Further, it was not investigated to what degree also different combinations of skewness of the predictors may adversely result in spurious interaction effects. To study these problems in more detail, we conducted a simulation study. Especially, we were interested in the question whether the inclusion of quadratic terms into the latent structural equation would reduce the occurrence of spurious interaction effects and make the inference about interaction more robust. When a latent predictor (e.g., \( \xi_1 \)) is skewed, its third centered moment is nonzero, which implies that the squared predictor (e.g., \( \xi_1^2 \)), or even the product term (e.g., \( \xi_1 \xi_2 \)) in case of correlated predictors, can be correlated with the linear predictor (e.g., \( \xi_1 \)). Thus, it is not surprising that a situation with skewed predictors in the population may lead to spurious interaction or quadratic effects when nonlinear terms are added to the linear model equation.

For the simulation study, data were generated under three different conditions for the skewness of the latent predictor variables: Condition I with normally distributed predictors, condition II with opposite sign of skewness for the two predictors, and condition III with similar sign of skewness for the two predictors. The data generation feature of the EQS program (Bentler, 1995) was used, which allows the specification of values for skewness and kurtosis for the variables. Only the data for the predictors were generated as nonnormal, whereas the residual and error variables' data were always generated as normal. After the data for the exogenous variables were all generated in EQS, the data for the indicator variables were computed from these according to the model equations, using a Pascal program. For the study, a basic linear structural equation without interaction term was selected for data generation:
η = α + γ₁ξ₁ + γ₂ξ₂ + ζ, \hspace{1cm} (7)

where η is the latent criterion, α is the intercept, γ₁ and γ₂ are structural coefficients, ξ₁ and ξ₂ are latent predictors, and ζ is the disturbance term. The following parameter values were selected: α = 1.00, γ₁ = 0.20, γ₂ = 0.40, φ₁₁ = 0.49, φ₂₁ = 0.235, φ₂₂ = 0.64, and ψ = 0.20. The correlation among the predictors was 0.42. Each predictor was measured by two indicators, with one of the two loadings scaled to 1.0 for each predictor. The loadings of the unscaled indicators were \( \hat{\lambda}_{x11} = 0.60 \) and \( \hat{\lambda}_{x42} = 0.70 \). The predictor ξ₁ was measured with reliabilities .49, .22, and ξ₂ was measured with reliabilities .64, .38. The criterion variable η was measured by one y-variable without error: y = η. The reliabilities were selected to be not very high in order to evaluate the performance of the method under reasonably difficult conditions. The specified model has 13 parameters and 5 indicator variables. For condition II, the skewness/kurtosis of ξ₁ was -2.0/6.0 and the skewness/kurtosis of ξ₂ was +1.5/5.0; for condition III, the skewness/kurtosis of ξ₁ was +2.0/6.0 and the skewness/kurtosis of ξ₂ was +1.5/5.0. 200 data sets of sample size N = 400 for the five indicator variables were generated under each of the three conditions. The data sets were then analyzed with LMS using Mplus 3.1. Under each condition, the data were generated based on a linear model (Equation (7)) and then analyzed with two different models: With Model A as a Kenny-Judd type interaction model (Kenny & Judd, 1984) where a latent product term \( \omega_{12} \xi_1 \xi_2 \) was added to Equation (7), and with Model B where both a latent product term \( \omega_{12} \xi_1 \xi_2 \) and two latent quadratic terms \( \omega_{11} \xi_1^2 \) and \( \omega_{22} \xi_2^2 \) were added to Equation (7). In total, this resulted in a 3×2 study design. Table 1 shows the estimation results. We restrict our report to the coefficients of the product and quadratic terms, that is, to \( \omega_{12} \), \( \omega_{11} \), and \( \omega_{22} \). The estimates for the coefficients of the linear effects, γ₁ and γ₂, are not reported here; there was no sign of any considerable bias for those.

Although the scope of our simulation study was limited, it revealed a pattern that was also observed in other simulation studies not reported here. Under condition I, where the data for the predictors were generated according to a normal distribution, neither the analysis with Model A nor with Model B showed signs of inflated Type I error for the estimates of \( \omega_{12} \): For Model A, the Type I error rate was below the nominal rate of 5%; for Model B, the Type I error rate of 6.5% was only slightly above the nominal rate of 5%, but for a simulation study with only 200 replications such a deviation is not unusual. Under condition II, the data for the predictors were generated with opposite signs of skewness. Again, neither for Model A nor for Model B there were signs of seriously inflated Type I error for the estimates of \( \omega_{12} \). However, the Type I error rate found for \( \omega_{12} \) here was with 6.5% lower than the somewhat inflated rate of 10.4% found by Klein and Moosbrugger (2000) in an earlier study. As expected, under condition II the skewed predictors resulted in estimates of nonzero quadratic effects in some replications: The percentages of significant quadratic effects were 10.5% and 23% for \( \xi_1^2 \) and \( \xi_2^2 \), respectively.
Table 1. LMS/Plus estimation results under 3 conditions (normal predictors, predictors with opposite sign skewness, and predictors with same sign skewness). Model A included a latent product term, Model B included both latent product and quadratic terms. Displayed are means (M), standard deviations (SD), 95% coverage, and percentage of significant t-values of estimates. For the t-values, the nominal Type I-error rate was set at 5%.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Condition I (Normal), Model A (Latent Product Term)</th>
<th>Condition I (Normal), Model B (Latent Product and Quadratic Terms)</th>
<th>Condition II (Opposite Sign Skewness), Model A (Latent Product Term)</th>
<th>Condition II (Opposite Sign Skewness), Model B (Latent Product and Quadratic Terms)</th>
<th>Condition III (Same Sign Skewness), Model A (Latent Product Term)</th>
<th>Condition III (Same Sign Skewness), Model B (Latent Product and Quadratic Terms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ω_{12}</td>
<td>ω_{11}</td>
<td>ω_{22}</td>
<td>ω_{12}</td>
<td>ω_{11}</td>
<td>ω_{22}</td>
</tr>
<tr>
<td></td>
<td>0.007</td>
<td>0.046</td>
<td>97.5</td>
<td>0.012</td>
<td>0.132</td>
<td>94.5</td>
</tr>
<tr>
<td></td>
<td>2.5%</td>
<td></td>
<td></td>
<td>6.5%</td>
<td>90.5%</td>
<td>10.5%</td>
</tr>
<tr>
<td></td>
<td>93.0</td>
<td>95.5</td>
<td>77.0</td>
<td>98.0</td>
<td>77.0</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td>95.0</td>
<td></td>
<td></td>
<td>51.5%</td>
<td>92.0</td>
<td>8.0%</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
<td></td>
<td></td>
<td></td>
<td>6.0%</td>
<td>23.5%</td>
</tr>
<tr>
<td></td>
<td>2.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Under condition III, the data for the predictors were generated with similar signs of skewness. The results in Table 1 show that this condition is the most critical in our study. First, not surprisingly for condition III, the skewed predictors resulted in a bias for the quadratic effect parameters ω_{11} and ω_{22} in Model B. When the data were analyzed with a standard Kenny-Judd type of interaction model (Model A), spurious interaction effects now posed a serious problem: In 48.5% of the replications, the t-value for the interaction parameter ω_{12} was significant when tested at a nominal 5% Type I error rate. Also, Model A showed a clear bias for the estimates of ω_{12}, and 95% coverage was critically low with 51.5%. This demonstrates very clearly that the Kenny-Judd type model would result in spurious interaction as a consequence of variable skewness, which shows the severe limitations of Model A.
In sharp contrast to this, when the data were analyzed with a baseline model that included quadratic terms $\xi_1^2$ and $\xi_2^2$ (Model B), the Type I error rate dropped to a very acceptable rate of 8%. Although the quadratic effects were not significant in the majority of replications under condition III, they effectively served the purpose of correctly adjusting the effect of the product term. In summary, the results of the simulation study suggest:

1. There is strong evidence that by using quadratic terms in the latent baseline model, the problem of spurious latent interaction effects in structural equation modeling can be effectively controlled and critically inflated Type I error rates for the interaction parameter $\omega_{12}$ can be avoided.

2. The parameter estimates for the quadratic terms in the baseline model can be affected by the skewness of the predictors; therefore, when strong skewness is detected among the indicators of the latent predictor variables, the estimates for the coefficients of latent quadratic terms should be interpreted with caution and may not represent an actual quadratic relationship between criterion and predictors. Nevertheless, the quadratic terms may be kept in the model equation, as they effectively serve the purpose of adjusting the estimate of the interaction parameter.

EMPIRICAL EXAMPLE

We present an empirical example of an interaction effect analysis for behavioral data using the LMS method that has been implemented in Mplus. The data have been drawn from the Oregon Youth Study (OYS), where a grade 4 cohort of boys and their families has been recruited from schools in the Eugene-Springfield metropolitan area of Oregon (US) serving areas characterized by higher rates of delinquency. Previous analysis of data drawn from this study has been described by Klein and Stoolmiller (2003). The sample was primarily a working class (50%), European American (86%), two-parent sample (70%). The purpose of the interaction model analysis was to investigate in what ways children exert evocative effects on parental discipline (Vuchinich, Bank & Patterson, 1992). We examined the moderator hypothesis that parental perceptions (measured at boy age 10) of early child manageability problems in the first five years of the boy’s life, moderate parental discipline responses to the child’s disruptive behavior at age 10. It was specifically hypothesized that the more parents have perceived their boys to be difficult to manage even from early on in life, the more aversive will now be their reactions to his current disruptive behavior. We further expected that the regressive relationship between boy’s disruptive behavior (as independent variable) to evoke parental harsh discipline (as dependent variable) is positively moderated by the parent’s perception of early child manageability problems (as moderator variable).

The original sample consisted of $N = 206$ observations. Among the variables measured were the four variables: early manageability problems ($x_1$), proportion measure of boy’s disruptive behavior ($x_2$), rate measure of boy’s disruptive behavior ($x_3$), and microsocial measure of aversive behavior directed to the boy ($y$). The data were initially screened for missing values, outliers, and deviations from
the normal distribution. 23 cases were deleted because they had either missing values or outliers, resulting in $N = 183$ cases. Skewness and kurtosis were checked for the x-variables, because the LMS method was developed under the assumption that the latent predictors and the errors are normally distributed. The x-variables displayed some skewness. The skewness of the variables $x_1$, $x_2$, and $x_3$ was 0.45, 1.33 and 1.12, respectively; the kurtosis of $x_1$, $x_2$, and $x_3$ was −0.92, 2.04, and 0.96, respectively.

The data were analyzed with Model A and Model B as specified above. Additionally, an additive quadratic model (Model C) was analyzed. The structural equations of the models are

$$
\eta = \alpha + \gamma_1 \xi_1 + \gamma_2 \xi_2 + \omega_{12} \xi_1 \xi_2 + \zeta \quad \text{(Model A)},
$$

$$
\eta = \alpha + \gamma_1 \xi_1 + \gamma_2 \xi_2 + \omega_{11} \xi_1^2 + \omega_{22} \xi_2^2 + \omega_{12} \xi_1 \xi_2 + \zeta \quad \text{(Model B)},
$$

$$
\eta = \alpha + \gamma_1 \xi_1 + \gamma_2 \xi_2 + \omega_{11} \xi_1^2 + \omega_{22} \xi_2^2 + \zeta \quad \text{(Model C).} \tag{8}
$$

The first two equations contain an interaction term $\omega_{12} \xi_1 \xi_2$ that models a possible interaction between the child’s early manageability problems and parental harsh discipline. The measurement of the latent variables is given by the measurement model

$$
x_1 = 1.0 \xi_1 , \quad x_2 = 1.0 \xi_2 + \delta_2 ,
$$

$$
x_3 = \lambda_{32} \xi_2 + \delta_3 , \quad y = 1.0 \eta . \tag{9}
$$

In our case, there is only one indicator variable, $x_1$, available for the measurement of the construct $\xi_1$ (boy’s disruptive behavior). Thus, $\xi_1$ is assumed to be measured without error by $x_1$. The measurement equations don’t include intercepts because standardized data were used for analysis. For both models, analysis was carried out in Mplus 3.1 that allows for the specification of product and quadratic terms composed of latent predictors. The estimation results with estimates and t-values for the structural parameters are given in Table 2. The estimated correlation between the two predictors was close to 0.07 across all models. Because of the weak correlation, it was expected that the inclusion of quadratic terms would not change the estimate for the interaction parameter $\omega_{12}$ very much for this particular data set.

First, we note that in Table 2 the estimates for $\gamma_1$ are not significant, and $\gamma_2$ is estimated about equally across all models. The estimate of the interaction parameter $\omega_{12}$ stays stable when one moves from Model A to Model B and adds the two quadratic terms to the model. This suggests that the interaction term $\omega_{12} \xi_1 \xi_2$ in Model A does not contain any substantial additive component that would be removed when quadratic terms are entered. None of the quadratic terms are significant in Model B or C. However, when one compares the estimates for the quadratic effect parameters $\omega_{11}$ and $\omega_{22}$ across models B and C, it is noted that the estimates for $\omega_{11}$ and $\omega_{22}$ are larger for Model C, where the interaction parameter is set to zero. In terms of explanation of variance of the criterion variable, the quadratic terms in
Table 2. LMS/Mplus estimation results for behavioral data. Three models were analyzed: Model A with a latent product term, Model B with both latent product and quadratic terms, and Model C with quadratic terms only. All models included linear terms for the predictors. Displayed are estimates and t-values for the structural parameters.

<table>
<thead>
<tr>
<th></th>
<th>Model A (product term)</th>
<th></th>
<th>Model B (product and quadratic terms)</th>
<th></th>
<th>Model C (quadratic terms)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>t</td>
<td>Estimate</td>
<td>t</td>
<td>Estimate</td>
<td>t</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.072</td>
<td>1.136</td>
<td>0.045</td>
<td>0.604</td>
<td>0.034</td>
<td>0.477</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>0.473</td>
<td>4.927</td>
<td>0.485</td>
<td>4.202</td>
<td>0.456</td>
<td>3.900</td>
</tr>
<tr>
<td>$\omega_{12}$</td>
<td>0.218</td>
<td>2.735</td>
<td>0.217</td>
<td>2.642</td>
<td>-/-</td>
<td>-/-</td>
</tr>
<tr>
<td>$\omega_{11}$</td>
<td>-/-</td>
<td>-/-</td>
<td>0.058</td>
<td>0.926</td>
<td>0.075</td>
<td>1.188</td>
</tr>
<tr>
<td>$\omega_{22}$</td>
<td>-/-</td>
<td>-/-</td>
<td>-0.006</td>
<td>-0.087</td>
<td>0.041</td>
<td>0.597</td>
</tr>
</tbody>
</table>

Model C compensates to a certain degree for the omitted product term. This parameter change affects the likelihood-ratio test of the interaction parameter. The likelihood-ratio test for testing Model B against Model C gives a value of $\chi^2 = 3.6$ (d.f. = 1) for the test of the interaction parameter, which is close to the cutoff value of 3.84, but not significant at the 5% Type I error level. The t-value for $\omega_{12}$ suggests it is significant ($t = 2.642$). But the likelihood-ratio statistic is more reliable here than the t-value, because the estimates of the quadratic effect parameters change from Model B to Model C, whereas a parameter test based on the t-value assumes stability of the remaining parameter estimates. Altogether, we conclude that there is some evidence of an interaction effect of child early manageability problems and child disruptive behavior on parental harsh discipline, although it is not significant. The use of the product term in Model B and the stability of the estimate for the interaction parameter across Models A and B additionally supports the interpretation that there may exist an interactive, non-additive relationship between the variables. However, perhaps a larger sample or the inclusion of additional indicators would be necessary here to further decide with greater certainty about the interaction effect hypothesis.

DISCUSSION

This paper addresses the practical question of how spurious interaction effects can be avoided in regression and latent variable analysis. In Sections 1 and 2, the common practice to use a product term to model interaction and the recommendation of some researchers to use quadratic terms to adjust for spurious interaction are critically revised from a theoretical standpoint. Interactive relationships are characterized as non-additive regression relationships, and it is shown that under certain linearity assumptions about the predictor variables, a baseline regression model with linear and quadratic terms as regressors is sufficient to avoid spurious interaction. In Section 3, spurious interaction in latent variable models is investigated. Supported by a simulation study, we demonstrate that when
linear and quadratic terms are used in a structural equation model that has a product term to model interaction, spurious interaction that is a consequence of skewness in the predictors can be kept under control. Type I error rates that are grossly inflated when a standard Kenny-Judd type interaction model is used are now effectively reduced to values close to their nominal rate. The usefulness of the LMS method, which has been implemented in Mplus, to carry out such analyses is demonstrated. These findings suggest that having quadratic terms in the baseline model appears to be very useful for the analysis of latent interaction effects in contexts where it is of critical importance whether an additive model should be abandoned in favor of an interaction model. Further studies, with different sizes of skewness, different sizes of interaction and quadratic effects, and different sample sizes, are requested to investigate in more detail under what conditions the use of quadratic terms can be safely recommended as a routine procedure in the analysis of latent interaction effects.

REFERENCES


ASSESSING SPURIOUS INTERACTION EFFECTS


*Andreas G. Klein*
*The University of Western Ontario*

*Karin Schermelleh-Engel*
*Helfried Moosbrugger*
*Augustin Kelava*
*Johann Wolfgang Goethe University Frankfurt*