Structural Equation Modeling and Latent Variable Approaches

ALEX LIU

Abstract

Structural equation modeling and latent variable approach (SEM) is experiencing rapid development with wide application as a result of using big data and modern computing technologies. This essay first gives an introduction of SEM, and then summarizes the foundational research in developing better fit indices and in developing more efficient computing algorithms. Also, we review two most important cutting-edge researches in using SEM for causal analysis and in managing workflows of SEM. For the future SEM research, we have discussed issues of big data, new applications, equivalent models, and hybrid modeling.

INTRODUCTION

Structural equation modeling and latent variable approach (SEM) is a set of modeling techniques used by researchers to represent, estimate, and interpret complicated relationship among many variables. SEM's model is often represented by a set of more than one regression type of equations. SEM is capable of modeling causal relationships, as well as direct and indirect impacts of many variables over other variables. SEM is also a model development process leading to a final SEM model with a best data fit. SEM has recently experienced rapid development with wide applications as a result of using big data and modern computing technologies.

SEM has been considered as an extension of regression modeling, but it is different from regression modeling as its coefficient fitting methods rely on aggregate data rather than individual data points. Specifically, in coefficient estimation, SEM minimizes the differences between data-generated covariance matrix and model-predicted covariance matrix, while regression modeling minimizes the differences between observed dependent variable values and predicted dependent variable values.

Structural equation models support both confirmatory and exploratory modeling. That is, SEM can be used for both theory testing and theory

Emerging Trends in the Social and Behavioral Sciences. Edited by Robert Scott and Stephen Kosslyn.
© 2015 John Wiley & Sons, Inc. ISBN 978-1-118-90077-2.

development, for the later approach, a step-by-step approach is often taken. The common characteristics of SEM model building processes may be summarized with 4Es—equation, estimation, evaluation, and explanation—as given.

EQUATION SPECIFICATION

Either used for a confirmatory hypothesis testing or exploratory model development, an initial model must be correctly specified as based on the theories or hypotheses that the researcher is attempting to confirm, or based on initial ideas for a model development plan.

Most SEM model consists of two main components of models: the structural model representing potential causal dependencies between endogenous and exogenous variables, and the measurement model representing relations between latent variables and their indicators.

To represent a specified model, researchers often use path diagrams, matrices, or tables.

ESTIMATION OF COEFFICIENTS

The next step after model specification is to estimate coefficients for all of the SEM equations, usually by minimizing the difference between the covariance matrices calculated from the data and the covariance matrices inferred by the SEM equations. This is often obtained through numerical maximization of a fit criterion by implementing maximum likelihood estimation algorithms. This is often accomplished by using a specialized SEM algorithm, provided by SAS, R, SPSS, and Mplus .

EVALUATION OF MODEL WITH MODEL FIT INDICES

Once all the coefficients are estimated, researchers move forward to examine the "fit" of the estimated SEM model. This is a basic task in SEM modeling that forms the base for accepting or rejecting models for confirmatory approach and developing model modification hints for exploratory approach.

Some of the commonly used measures of fit include the following:

- *Chi-Squared.* This is a basic measure of fit as used in the calculation of many other fit measures. In calculation, it is a function of the sample size and the difference between the observed covariance matrix and the model inferred covariance matrix.
- *Root Mean Square Residuals.* an index directly calculating the difference between the observed covariance matrix and the model inferred covariance matrix.

- Goodness-of-Fit statistic (GFI). Relative amount of the variance in the observed covariance matrix as predicted by the model inferred covariance matrix.

Model Modification—Going Back to Es. Model evaluation does not necessarily lead to accept or reject a model, but often leads to model modification, especially when a model development approach is taken. That is, this step often leads the researchers to go back to the equation specification or estimation step.

Explanation and Communication

Once researchers are satisfied with the model evaluation results, some final models are selected. Then, their coefficients need to be interpreted so that the relationship between measures and their indicators, and the causal relationships among all variables can be made clear.

Special caution should always be taken when making claims of causality, while total effects, direct effects, and indirect effects are often calculated and used for explanation.

FOUNDATIONAL RESEARCH

Developing better model fit indices and developing more efficient computing algorithm for coefficient estimation are considered as the most important foundational research.

Model Fit Indices

There are many model fit indices available to evaluate how an SEM model fits with a data, with more and more of these indices being developed. Among them, the most important three are absolute fit indices, incremental fit indices, and parsimonious fit indices.

Absolute Fit Indices. Absolute fit indices demonstrates directly which model has the best data fit, and are derived directly from the fit of the observed and predicted covariance matrices and the MLE minimization function. These fit measures provide the basic indication of how well the developed model fits the data. Included in this category are the chi-squared, RMSEA, GFI, AGFI, the RMR and the SRMR.

Model Chi-Square

Chi-square value is the most commonly used and the most basic measure for evaluating a model's overall model fit, and for testing the hypothesis about the discrepancy between the sample and fitted covariance matrices. A good model fit would provide an insignificant Chi-square test result at a 0.05 threshold; thus, the Chi-square statistic is often referred to as either a "badness of fit" or a "lack of fit" measure with higher value to indicate badness.

Chi-square statistics have some limitations. First, this statistical test assumes multivariate normality of the data so that any severe deviation from normality may result in model rejections even when the model fits the data well. Second, the approximation to the Chi-square distribution only occurs for large samples that this statistical significance test is sensitive to sample size. So in situations where small samples are used, this Chi-square statistic lacks power of discriminating between good fitting models and poor fitting models.

To overcome the restrictiveness of the Model Chi-Square, researchers have been continuing to create alternative indices of assessing model fit. One example of them is a statistic like Wheaton *et al.*'s relative/normed chi-square (χ^2/df) that tries to minimize the impact of sample size on the Model Chi-Square.

More and more indices are expected for this line of research.

Root Mean Square Error of Approximation (RMSEA)

The RMSEA was first developed by Steiger and Lind in the 1990s. The RMSEA measures the different between the model predicted covariance matrix and the covariance matrix calculated directly from data with a smaller number to indicate better fit.

On the basis of recent research, recommendations for RMSEA cut-off points have been reduced considerably in the past 15 years. Up until the early nineties, an RMSEA in the range of 0.05–0.10 was considered an indication of fair fit and values above 0.10 indicated poor fit. It was then thought that an RMSEA of between 0.08 and 0.10 provides a mediocre fit and below 0.08 shows a good fit. However, more recently, a cut-off value close to 0.06 or a stringent upper limit of 0.07 seems to be the general consensus amongst experts in this area. Effort has also been made for an RMSEA confidence interval to be calculated easily, per recent research.

Goodness-of-Fit Statistic (GFI) and the Adjusted Goodness-of-Fit Statistic (AGFI)

The GFI was created by Jöreskog and Sorbom initially as an alternative to the Chi-square test. GFI calculates the proportion of covariance variance that is accounted for by the predicted covariance, with a value ranging from 0 to 1. When there are a large number of degrees of freedom in comparison to sample size, the GFI has a downward bias. In addition, it has also been found that the GFI increases as the number of parameters increases and has an upward bias with large samples. Traditionally a cut-off point of 0.90 has been recommended for the GFI. However, simulation studies have shown that when factor loadings and sample sizes are low, a higher cut-off of 0.95 is more appropriate. Given the sensitivity of this index, it has become less popular in recent years. Related to the GFI is the AGFI which adjusts the GFI based upon degrees of freedom, with more saturated models reducing fit. Thus, more parsimonious models are preferred while penalized for complicated models.

Root Mean Square Residual (RMR) and Standardized Root Mean Square Residual (SRMR)

The RMR is the square root of the difference between the residuals of the sample covariance matrix and the model inferred covariance model, while SRMR is the standardized RMR. Both RMR and SRMR are the simplest fit indices, and also the most intuitive ones.

Further research is expected to understand more about using residual differences to measure model fit.

Incremental Fit Indices. Incremental fit indices, also known as comparative or relative fit indices, do not use fit statistics or fit measures in its raw form but compare fit statistics of the developed model against that of a baseline model.

Normed-Fit Index (NFI)

The most basic of these incremental fit indices is the normed fit index (NFI). This statistic assesses the model by comparing the χ^2 value of the model to the χ^2 of the null model. The null/independence model is considered as the worst case scenario as it specifies that all measured variables are uncorrelated. Values for this statistic range between 0 and 1 with values greater than 0.90 indicating a good fit. A major drawback to this index is that it is sensitive to sample size, underestimating fit for samples less than 200.

CFI (Comparative Fit Index)

The comparative fit index (CFI) is a revised form of the NFI that takes sample size into consideration, so that it performs well even when sample size is small. This index was first introduced by Bentler and subsequently included as part of the fit indices in many computing programs including LISREL. Similar to the NFI, this statistic assumes that all latent variables are uncorrelated that constitutes a null/independence model and then compares the sample covariance matrix with this null model. As with the NFI, values for this statistic range between 0.0 and 1.0 with values closer to 1.0 indicating good fit. A cut-off criterion of CFI ≥0.90 was initially advanced; however, recent studies have shown that a value greater than 0.90 is needed in order to ensure that mis-specified models are not accepted. From this, a value of CFI ≥0.95 is presently recognized as indicative of good fit.

Parsimony Fit Indices. Parsimony fit indices also known as parsimony-corrected fit indices are developed to overcome the problem of obtaining good fit indices at the cost of saturating a model, for which two types of parsimony of fit indices have been developed: the parsimony goodness-of-fit index (PGFI) and the parsimonious normed fit index (PNFI). The PGFI is based on the GFI by adjusting for degrees of freedom. The PNFI also adjusts for degrees of freedom; however, it is based on the NFI. That is, both of them penalize for degrees of freedom, so they result in fit values that take parsimony into consideration.

A second form of the parsimony fit index is those known as information criteria indices. The best known of these indices is the Akaike Information Criterion (AIC). These statistics are generally used when comparing nonnested or nonhierarchical models estimated with the same data and indicates to the researcher which of the models is the most parsimonious. Smaller values suggest a good fitting.

It is worth noting that most of the fit statistics need a sample size of 200 to make their use reliable. In addition, fit indices may point to a well-fitting model but in fact parts of the model can fit poorly. Therefore, research has been developed to measure fit of the whole model, as well as submodels.

A lot of research has been completed recently with good results, but considerable controversy has also come up concerning fit indices. Some researchers claim that fit indices do not add anything to the analysis and only the model chi square should be interpreted. The major concern is that fit indices may allow some researchers to claim a miss-specified model as a good model. Other researchers argue that cutoffs for a fit index can be misleading and subject to misuse, even all fit indices are good. Overall, most scholars believe in the value of fit indices, but caution against reliance on cutoffs, for which more research is expected.

Efficient Computing and Identification. One of the SEM key issues is about how to obtain the unbiased coefficients of some proposed SEM models efficiently. And, this issue deserves even more attention here, as SEM faces many special computing challenges that other modeling approaches do not have to.

First, not all the SEM models are identifiable, while no identification issues exist for other approaches such as regression modeling.

Second, not all the SEM algorithms can guarantee results. Some numerical iterative algorithms do not converge to return estimations for some models and data, such as the algorithms provided by the popular software LISREL.

MLE is the common used methods for now, as it is asymptotically unbiased and most efficient. However, the unbiasedness and efficiency may not hold for data deviated from non-normal distribution. Outlying cases or nonnormally distributed data can make the ML estimator (MLE) biased and inefficient. Many other estimation methods such as robust estimation methods have been developed.

Currently, most researchers use SPSS or SAS or LISREL or EQS for SEM computing. These four algorithm providers still play a leading role in SEM computing.

However, Mplus is worth noting, as it offers great innovation including some to solve identification issues and convergence issues. It also makes using categorical indicators and multilevel analysis easy, and even with Bayesian analysis incorporated into the software system.

Another good alternative is R. R is a popular open source statistical software used by millions of researchers. In R, some packages including SEM, LAVAAN, and OpenMX have been developed and started to be widely used. As R is an open source platform, more and more SEM packages are expected to come, with innovative algorithms being incorporated quickly.

CUTTING-EDGE RESEARCH

A lot of work has been done to further develop SEM. Here, we discuss two of the most significant cutting-edge researches that are to derive causal insights from SEM and to develop good workflows for SEM.

Causal Inferences

Many researchers come to SEM for causal analysis methods. Therefore, using SEM to infer causal relationship has big audience. Among all the work of developing causality out of SEMs, professor Judea Pearl's work is the leading one, as summarized in his seminal book CAUSALITY and led him to win the 2012 Tuning award.

The main purpose of professor Judea Pearl's book CAUSALITY is to

- develop graphical tools in representing and assisting causal analysis;
- discuss about causality out from SEMs;
- develop algorithms using partial correlations to discover causal structure under certain assumptions.

Experts of research methods often say that "research methods do not equal to statistics." Research methods equal statistics plus something else. Pearl's work is to formalize this "something else" and provide tools to work on them explicitly. In traditional empirical analysis, at least in the mainstream methods teaching, this "something else" for causal analysis is that variable A is a cause of variable B, if:

- A and B are correlated.
- The association arises because A causes B and not vice versa because of temporal or logical or theoretical reasons.
- The association between A and B is not spurious.

According to Pearl, statistics deals with mean, variance, correlation, regression, dependence, conditional independence, association, likelihood, collapsibility, risk ratio, odd ratio, marginalization, conditionalization, "controlling for," … while causal analysis deals with randomization, influence, effect, confounding, "holding constant," disturbance, spurious correlation, instrumental variables, intervention, explanation, attribution, … The second part minus the first part is the "something else."

Professor Pearl's language to formally represent causal analysis and its components include both structural equation models (linear, nonlinear, and nonparametric) and graphical diagrams. Pearl uses do(x) to represent intervention. As many methodologists will agree, with Pearl's work, method concepts such as spuriousness and confounding are much better formalized than ever before.

His proposed rules of causal analysis include criterion to select covariates for adjustment, intervention calculus, and counterfactual analysis. Professor Pearl also proposed IC* algorithm to discover causal structures.

These are good contributions made by Pearl's work. But, this is just a beginning. In general, there are more questions than answers in his book CAUSALITY. There are also many missing links we need to bridge, in order to conduct a good causal analysis. For example, indirect effects are not covered as much as the direct effects and total effects. How to estimate the strength of a causal influence is also left out.

Professor D.A. Freedman of UC Berkeley takes a different view than that of Pearl (Freedman, 2004). Freedman claims that Pearl's work is based on many assumptions that are unrealistic and difficulty to confirm in applied research. In other words, Pearl's causal analysis needs both SEM work and some assumptions.

Published in 1993 (2nd edition in 2000 by MIT Press), the book Causation, Prediction and Search summarize another line of work by Spirtes, Glymour, and Scheines (SGS) of Carnegie Mellon University that is more data driven, and where they actually developed a software for their developed algorithms and applied to a lot of real research.

In general, to successfully infer causality from statistical evidence such as correlation does require some subject knowledge, additional statistical methods, and hard work. However, the work of Pearl and SGS helped to improve the current practice greatly.

Initial Model Generation

Specifying an initial model is a key element and the first step for any SEM modeling. In the past, initial model is often derived from qualitative research or from existing theories. Currently, many researchers claim initial model could be generated from data and even automatically, for which quite a few outstanding research projects are merging out. The IC algorithm proposed by professor Judea Pearl and the tetrad software developed by the SGS group of Carnegie Mellon University are among the good examples in this area.

Per TETRAD project introduction, TETRAD is a program that creates, simulates data from, estimates, tests, predicts with, and searches for causal and statistical models. The program provides sophisticated methods in a friendly interface requiring very little statistical sophistication of the user and no programming knowledge. Tetrad is a freeware.

Tetrad is unique in the suite of principled search ("exploration," "discovery") algorithms it provides—for example, its ability to search when there may be unobserved confounders of measured variables, to search for models of latent structure, and to search for linear feedback models—and in the ability to calculate predictions of the effects of interventions or experiments based on a model. All of its search procedures are "pointwise consistent"—they are guaranteed to converge almost certainly to correct information about the true structure in the large sample limit, provided that structure and the sample data satisfy various commonly made (but not always true!) assumptions.

Tetrad is limited to models of categorical data (which can also be used for ordinal data) and to linear models ("structural equation models") with a Normal probability distribution, and to a very limited class of time-series models. The Tetrad programs describe causal models in three distinct parts or stages: a picture, representing a directed graph specifying hypothetical causal relations among the variables; a specification of the family of probability distributions and kinds of parameters associated with the graphical model; and a specification of the numerical values of those parameters.

The program and its search algorithms have been developed over several years, and are currently being used by researchers and practitioners.

Per tests performed by many researchers, both the algorithms developed by Pearl and SGS do not work as well as claimed by their authors, even they do give good hints to develop initial models. Professor Freedman of UC Berkeley claims these algorithms do not work at all, because they are based on false assumptions. As I know, quite many scholars including myself tried these algorithms on some empirical data, and found these algorithms often lead us to nowhere or to some errors. However, many ideas presented in these algorithms can be used, in combination with subject knowledge and other statistical methods such as SEM method, to aid us in generating hypotheses and also in generating initial models. Professor Bill Shipley of Universite de Sherbrooke has some good work along this line.

WORKFLOW MANAGEMENT

To many researchers, SEM is also a model development workflow leading to a final SEM model with a best data fit. That is, good SEM always involves a long workflow, so that a good workflow management system could take SEM to a higher level.

Given here is an example of a common used workflow of building a SEM model from a sampling survey data.

- Check data structure to ensure a good understanding of the data
  - Is the data a cross-sectional data? Is implicit timing incorporated?
  - Are categorical variables used?
- Check sampling
  - What is the population and sampling method?
- Check missing values
  - *don't know* or *forget* as an answer may be recoded as neutral
- OR treated as a special category
- some variables may have a lot of missing values
  - to recode some variables as needed
- Study qualitative background
- Use this study to form some hypotheses
- to select some key dependent variables as the effects
- Conduct some descriptive studies to begin telling stories
  - use comparing means and cross-tabulations
  - check variability of some key variables (st dev and variance)
- Select groups of independent variables (exogenous variables)
  - as candidates of causes

Basic descriptive statistics

mean, st dev and frequencies for ALL variables

Measurement work

study dimensions of some measurements (EFA exploratory factor analysis may be useful here)

may form measurement models

Local models

Identify sections out from the whole picture

To explore about relationship

use cross-tabulations

graphical plots

use logistic regression

use linear regression

Conduct some partial correlation analysis to help model specification

Propose structural equation models by using results of (9)

identify main structures and sub structures

Connect measurements with structure models

Initial fits

create data sets for SEM software

programming in SEM software

Model modification

Use SEM results (mainly model fit indices) to guide

Reanalyze partial correlations

Diagnostics

distribution

residuals

curves

Final model estimation may be reached here

if not repeat step 13 and 14

Explaining the model (causal effects identified and quantified)

Obtaining good statistical models is usually not the end of a research. Deriving causal inference or causal explanation is often needed in order to make the research useful.

In this step, as a kind of common practice, theoretical knowledge and logic and common sense are frequently used, but often in an implicit and informal way.

In recent years, many computing infrastructures have been developed to manage workflows, for various disciplines such as neuroscience and market science. With more and more support coming out from the computer science side, the research and development of SEM workflows are progressing parallel with the development of new workflow computing systems.

KEY ISSUES FOR FUTURE RESEARCH

There are many issues that could rapidly develop SEM further, if handled well. Here, we discuss some of these key issues.

BIG DATA

We are in the era of big data. According to IBM and other research, very day, we create 2.5 quintillion bytes of data—so much that 90% of the data in the world currently has been created in the past 2 years alone. This data comes from everywhere: sensors used to gather information, posts to social media sites, purchase transaction records, and cell phone signals to name a few. All this data is **big data. And now,** big data is the next frontier for innovation, competition, and productivity, for which SEM has a lot to contribute.

SEM will offer great solutions to model big data. First, big data often comes with hundreds or thousands of variables, for which, SEM is often needed to represent the complicated relationship among all the key variables.

Second, with so many variables in hands, one way is to treat them as indicators of a latent variable so to simplify and also to make it meaningful.

On the output side, SEM makes results easy to be interpreted and to be used, as SEM and latent variable approach will allow one to use big data to measure concepts, and use big data to infer total effects.

To deal with the big data, machine learning is a fast growing discipline, for which SEM can contribute solutions at least to feature selection and model stability issues.

NEW APPLICATIONS

SEM has been used in many fields in social studies like in education research and in consumer behavior research.

SEM could be expanded to many practical areas such as credit risk and program evaluation.

In credit risk, it has been recognized that every piece of data is a credit data so that SEM is in demand.

In program evaluation, causal analysis is in demand to understand the complicated relationship among factors and outputs, and especially to estimate the total effects of intervention.

Latent variable approaches with new applications are generating fundamental impacts in how people are viewing measurements. In recent years, many rankings have been developed to rank various subjects including education and democracy. However, each ranking has its own biases. SEM and latent variables approaches provide perfect tools to evaluate these biases, and help users to form best measures out from these rankings, which have been recognized.

As these practical areas are well funded, which may fuel a new life into SEM?

Equivalent Models

In SEM, many models are equivalent in that they fit the data almost the same, if measured by fit indices. A lot of research has been completed to attack this issue. Some of them established necessary and sufficient condition for equivalence of structural equation models. Many rules have been produced for the equivalent model generation.

However, SEM practitioners are still puzzled by how many equivalent models they can fit to a data, and if some complicated and simple models are equivalent. And especially, not many algorithms have been developed yet for researchers to choose among equivalent models.

A breakthrough is expected when equivalent model research can be combined with SEM workflow management, and with efficient computing.

Hybrid Approaches

Graphical modeling and Bayesian modeling could complement SEM greatly, as well as some other computing intensive methods. But so far, they are progressing separately.

A great breakthrough is expected in the interaction of SEM and other related approaches including graphical modeling and Bayesian approaches. Some Bayesian and SEM hybrid models are starting to merge.

Workflow management system may serve as a platform for further developing hybrid approaches, especially in dealing with big data.

REFERENCE

Freedman D. A. (2004). Statistical Models for Causation, UCB Statistics Technical Report No. 651. www.stat.berkeley.edu/-census/651.pdf

FURTHER READING

Bollen, K. A., & Pearl, J. (2013). Eight myths about causality and structural equation models. In S. L. Morgan (Ed.), *Handbook of causal analysis for social research*. Ithaca, NY: Springer.

Liu, A. (2009). *Building structural equation models in social science*. Los Angeles, CA: RM.

OpenMx. R for advanced structural equation modeling. Retrieved from http://openmx.psyc.virginia.edu/

Pearl, J. (2009). *Causality*. Cambridge, England: Cambridge University Press.

Shipley, B. (2000). *Cause and correlation in biology*. Cambridge, England: Cambridge University Press.

ALEX LIU SHORT BIOGRAPHY

Dr. **Alex Liu** is an expert of quantitative research methods and a director of the RMA, where he has provided analytics consultation to many large organizations such as AOL, the United Nations, US Government, IBM, Indymac Bank, Farmers Insurance, Scripps Networks and Ingram Micro. In the past, Mr. Liu worked as a research fellow for the Asia/Pacific Research Center at Stanford, as an adjunct assistant professor for the Marshall School of Business at USC, as a senior scientist for IBM research, as a senior consultant for the Beyster Institute of UC San Diego, and as a senior consultant for the Global Entrepreneurship Monitoring of the Babson College and the London Business School. From 2003 to 2011, Dr. Liu taught structural equation modeling to PhD candidates in the Paul Merage School of Business of the University of California at Irvine. Alex has a PhD of Sociology and a MS of Statistical Computing from Stanford University.

RELATED ESSAYS

Social Epigenetics: Incorporating Epigenetic Effects as Social Cause and Consequence *(Sociology)*, Douglas L. Anderton and Kathleen F. Arcaro

To Flop Is Human: Inventing Better Scientific Approaches to Anticipating Failure *(Methods)*, Robert Boruch and Alan Ruby

Repeated Cross-Sections in Survey Data *(Methods)*, Henry E. Brady and Richard Johnston

Ambulatory Assessment: Methods for Studying Everyday Life *(Methods)*, Tamlin S. Conner and Matthias R. Mehl

The Evidence-Based Practice Movement *(Sociology)*, Edward W. Gondolf

The Use of Geophysical Survey in Archaeology *(Methods)*, Timothy J. Horsley

Network Research Experiments *(Methods)*, Allen L. Linton and Betsy Sinclair
Longitudinal Data Analysis *(Methods)*, Todd D. Little *et al*.
Data Mining *(Methods)*, Gregg R. Murray and Anthony Scime
Remote Sensing with Satellite Technology *(Archaeology)*, Sarah Parcak
Quasi-Experiments *(Methods)*, Charles S. Reichard
Digital Methods for Web Research *(Methods)*, Richard Rogers
Virtual Worlds as Laboratories *(Methods)*, Travis L. Ross *et al*.
Modeling Life Course Structure: The Triple Helix *(Sociology)*, Tom Schuller
Content Analysis *(Methods)*, Steven E. Stemler
Person-Centered Analysis *(Methods)*, Alexander von Eye and Wolfgang Wiedermann
Translational Sociology *(Sociology)*, Elaine Wethington