

# A Systematic Review of Structural Equation Modelling in Social Work Research

Baorong Guo, Brian E. Perron and David F. Gillespie

---

*Baorong Guo is in the social work faculty at the University of Missouri–St Louis and she is a faculty associate at the Center for Social Development (CSD) at Washington University in St Louis. Her current research focuses on inclusion of the poor in asset-building, human service non-profits and socio-economic development. Her recent research projects include the effects of household assets on material hardships, funding streams for human service non-profits and evaluation research on asset-based policies in the USA and China. Brian Perron studies services for persons with serious mental illnesses and substance use disorders, with an emphasis on comorbidities. Dr Perron has a range of teaching interests, with his primary interests relating to his practice experience in community mental health. His training in services research was supported by fellowships from the National Institute of Mental Health (T32) and the National Institute of Drug Abuse (F31). Other areas of research/scholarly interest: treatment processes, service access and utilization, evidence-based practices. Professor Gillespie is a national expert in the area of disaster preparedness and mitigation, specifically the integration of social work and social services with emergency management. Professor Gillespie teaches social measurement and structural equation modelling at the doctoral level and organizational dynamics and disaster preparedness at the master's level. Professor Gillespie is a member of the FOCUS St Louis Regional Disaster Preparedness Task Force—a cross-section of 30 citizen leaders representing a range of perspective and skills charged with assessing the region's readiness for a major disaster*

Correspondence to Baorong Guo, Ph.D., University of Missouri-St Louis, School of Social Work, 590 Lucas Hall, One University Boulevard, St Louis, MO 63121, USA. Email: [guob@umsl.edu](mailto:guob@umsl.edu)

## Summary

Structural equation modelling (SEM) is serving an increasingly important role in developing knowledge for the social work profession. Numerous advances have made the software more user-friendly, enabling users to conduct analyses without fully understanding the underlying assumptions or the implications from their analytic decisions. Unlike other fields, there have not been any published reviews in social work research that systematically describe and critique the use of SEM. This study systematically reviewed how SEM is used in social work research and the extent to which it reflects best practices. Thirty-two articles from top-ranked social work journals published from 2001 to 2007 were examined. Of the different types of SEM, the most commonly used

was confirmatory factor analysis. Strengths of the research reviewed included examining models and measures not previously tested empirically and generating new insights into old topics through the use of SEM. Weaknesses included significant model modifications without theoretical justification or substantive interpretations. Suggestions are offered for improving applications of SEM in social work research. Specifically, we encourage social work researchers to test competing models, to make model modifications only when theoretically justified, to detail the process of making modifications and to use estimation procedures appropriate to the data.

**Keywords:** Structural equation modelling, systematic review, social work research

## Introduction

As a general statistical methodology that subsumes and extends correlation, regression, factor analysis and path analysis (Schumacker and Lomax, 1996), structural equation modelling (SEM) represents an important advancement in social work research. SEM is focused on testing causal processes inherent in our theories. Before SEM, measurement error was assessed separately and not explicitly included in tests of theory. This separation has been an obstacle to improving theory in social work research. With SEM, measurement error is estimated and theoretical parameters are adjusted accordingly—that is, measurement error is subtracted from parameter estimates. Thus, SEM is a fundamental advancement in social work theory construction because it integrates measurement with substantive theory.

Structural equation modelling is serving an increasingly important role in developing knowledge for the social work profession. Over ten years ago, efforts were made to introduce this statistical methodology to social work research along with recommendations for making it part of social work doctoral education (Gillespie *et al.*, 1995; Hamilton and Orme, 1990; Orme and Fickling, 1992; Orme, 1993, 1995). Today, SEM is becoming more widely used in social work research. With the availability of user-friendly statistical software, such as LISREL, AMOS and EQS, SEM has become more widely used and reported in social work journals.

Other fields have conducted systematic reviews to gauge the quality and quantity of the use of SEM (e.g. Breckler, 1990; DiStefano and Hess, 2005; MacCallum and Austin, 2000; Martens, 2005; Russell, 2002). Taking stock of research in a particular area of research is necessary to ensure a high standard of quality research and publications. In a review of SEM in counselling psychology research, Martens (2005) found that counselling psychology researchers were not applying SEM that aligned with various best practices recommended by SEM experts. In a review of factor analysis as reported in the *Personality and Social Psychology Bulletin*, Russell (2002) investigated the use of factor analysis over time and highlighted many

appropriate uses and abuses. To date, no reviews of SEM in social work research have been conducted and published, which is a significant gap, as it relates to ensuring quality research in the profession.

Concern about journal quality in social work research has been growing (Sellers *et al.*, 2004, 2006; SSWR, 2005). It has been suggested that social work journals often do not have knowledgeable reviewers with sufficient scholarly expertise to provide rigorous critiques (Jenson, 2005). Structural equation modelling research may be especially vulnerable to poor reviews, as SEM is an advanced set of statistical procedures that is not yet part of the core statistical curriculum of social work doctoral education. In addition, numerous advances have made the software more user-friendly, enabling users to have conducted an analysis without fully understanding the underlying assumptions or the implications from their analytic decisions (see Tomarken and Waller, 2005). Each SEM analysis involves numerous decisions, creating opportunities for errors to be made.

The purpose of this study was to systematically review the use of SEM in core social work journals in order to understand how it is used and the extent to which it accords with best practices. Specifically, two research questions are explored. First, how is SEM used in social work research? Second, to what extent does the use of SEM in social work research reflect best practices? Systematic reviews use pre-planned methods to retrieve and summarize characteristics for a set of original studies. An obvious advantage of conducting systematic reviews is that they can reduce biases and random errors in reviewing current literature.

## Method

### Retrieval and selection

The first stage of this study involved retrieving SEM research in social work. Given the inter-disciplinary nature of the field, social work researchers publish in health and social science journals as well as journals specific to social work. Rather than attempting to sample from all the journals, this systematic review targeted the top ten-ranked journals that were considered to publish the highest-quality social work research based on a national survey of social work faculty in the USA (Sellers *et al.*, 2006). These ten journals include *Social Service Review*, *Social Work Research*, *Journal of Social Service Research*, *Child Abuse and Neglect*, *Research on Social Work Practice*, *Journal of Social Policy*, *American Journal of Community Psychology*, *Children and Youth Services Review*, *Families in Society* and *Journal of Community Psychology*. Although these journals are mostly American-based, they are all international in orientation, publishing work from around the world. We believe these journals provide a good representation of SEM studies in social work. Our search included

studies published from January 2001 to January 2007. This six-year period yielded a sample size large enough to represent the range of SEM applications currently being published in social work. Applications of SEM in social work before 2001 were infrequent.

The initial retrieval involved an electronic search of journal article titles, keywords and abstracts in major databases (e.g. Academic Search Premier) with one or more of the above journals included. Specifically, abstracts and full text were selected when articles with keywords such as structural equation modelling (or structural equation models), path analysis, confirmatory factor analysis and latent growth curve appeared in the title. However, after numerous attempts, no set of keywords were found to be able to capture all relevant articles in the target journals. Thus, a hand search of each journal was conducted. The search resulted in a sample with fifty-eight studies that used any of the core SEM techniques (confirmatory factor analysis, path analysis, latent growth curves and structural equation models).

The selection involved two steps. The first step was to examine whether the retrieved articles were appropriate for addressing the research objectives of the current study. This examination resulted in the exclusion of one study because it was a methodological paper describing the use of SEM rather than applying SEM to test a theoretical model. The second step was to select SEM studies conducted by social work researchers. As this systematic review was centrally concerned with the quality of SEM research in social work, it was necessary to exclude studies that were not carried out by social work researchers. All author affiliations were examined and studies were excluded if neither the first nor the second author had an affiliation with a department or school of social work, social welfare or social service administration at the time of publication. In other words, studies that were authored by researchers in a department of psychology, sociology or other social science departments were excluded. Overall, twenty-five studies were excluded using this criterion. The exclusions resulted in an effective sample size of  $N = 32$ .

## Coding system

Since this study used the method of systematic review to assess selected SEM studies, a coding system had to be created to structure the review process. An initial coding strategy was developed, focusing on the type of SEM model analysed, software used, sample size, the testing of alternative models, theoretical justifications, multivariate normality, model modifications and model fit. This coding system was then pilot tested on a set of SEM articles not included in the final sample. To ensure validity of this coding system, an independent researcher with

an extensive record of SEM methodological papers reviewed all the coding items. This review led to slight revisions, which entailed additional pilot testing. The revised coding system was again tested by two researchers using a few SEM articles not in the final sample. After a 100 per cent agreement rate was achieved, a customized spreadsheet was developed and used to carry out the evaluation process. Table 1 summarizes the coding system used by this study.

## Coding and analytic strategy

Each of the articles in the sample was coded using the coding system created for this study. Preliminary coding was done by the first author of this study and double-checked by the other two authors to make sure that coding was correct. For a few of the articles with items that could

**Table 1** Coding system

Category	Item	Label	Values
Study identification	Author(s)	Last name of first author	String
	Year	Year of publication	Year
	Journal	Journal title	String
Study exclusions	SEM study	Did the study employ SEM?	1 = Include 2 = Exclude
	Social work faculty	Was the first or second author a social work faculty member?	1 = Include/social work faculty 2 = Exclude/non-social work faculty 3 = Indeterminable
Study type	SEM type	Classification of how SEM is used	1 = Confirmatory factor analysis 2 = Path analysis 3 = SEM* 4 = Latent growth curve
	Software	What software package was used?	1 = AMOS  2 = LISREL 3 = Mplus 4 = SAS (Proc Calis) 5 = EQS 6 = Mx 7 = Other 8 = Not reported
	Group	Did the study utilize a single or multiple group (stacked model) analysis?	1 = Single group  2 = Multiple group
	Recursivity	Did this study include any non-recursive (feedback) paths?	1 = Yes  2 = No

*Continued*

Table 1 Continued

Category	Item	Label	Values
Data preparation	Sample size	What is the sample size for the study?	Integer
	Random sample	Type of sampling strategy used	1 = Probability 2 = Non-probability 3 = Indeterminable
	Multivariate normality	Did the authors test for multivariate normality?	1 = Yes 2 = Not reported
	Multivariate outcome	If authors tested for multivariate normality, what was the outcome (normal, non-normal)?	1 = Normal 2 = Non-normal 3 = NA (not reported)
Specification	Number of correlated error terms	Number of correlated error terms in the final model	Integer
	Correlated error via modification	Were any correlated error terms added after the model was estimated—that is, as part of a modification process?	1 = Yes 2 = No 3 = Indeterminable
	Rationale for correlated error	What were the substantive reasons/interpretations of the correlated error terms?	String
	Number of latent variables	Number of latent variables included in the final model	Integer
	Number of observed variables	Total number of observed variables included in the final model	Integer
	Number of observed per latent variable	Number of observed variables per latent variable in the final model	Integer
	Estimation	Estimation method	What was the method of estimation?
Dichotomous/ordinal		Were any measures dichotomous or ordinal ( $\leq 4$ response categories)?	1 = Yes 2 = No 3 = Unclear
Model modification	Modification	Was the model modified for any reason? This includes all modifications except adding correlating error terms	1 = Yes

			2 = No 3 = Unclear Integer
	Modification: items dropped	How many items were dropped in the modification process?	
	Modification process	Extract a direct quote regarding the decision rules for the modification	String
	Modification substantive interpretations	What substantive or theoretical interpretations were made following the modification of the model?	String
Model fit	$\chi^2$ test	$p$ -value of $\chi^2$ test of final model	$p$ -value
	$\chi^2/df$ ratio	$\chi^2/df$ ratio of final model only if reported by authors	$\chi^2/df$ ratio
	Fit indices	Specify the specific indices and the values for each index	Value of indices
	Loadings and coefficients	Did authors report all factor loadings/path coefficients of final model?	1 = Yes
	Error variances	Did authors report all error variances of the final model?	2 = No 1 = Yes
	Modification indices	Did authors reference modification index in assessment of model fit?	2 = No 1 = Yes
	Standardized residuals	Did authors reference standardized residuals in assessment of model fit?	2 = No 1 = Yes
Reporting	Replication	What type of information was provided to allow replication of the study?	2 = No 1 = Covariance matrix  2 = Correlation matrix 3 = Available from author 4 = None
	Author's assessment of final model	What was the final assessment of the model based on the perspective of the author?	1 = Good fit
	Coder's assessment of final model	Comments by the coder reflecting the quality of the methodology employed	2 = Good fit with some problems, but fit is 'good enough' to move forward 3 = Poor fit 4 = Inconclusive String

---

\* These are models with both measurement models and structural relations.

not be decided, the three authors discussed each item to reach an agreement on its coding.

The assessment was done by summarizing descriptive statistics for all coded items. These descriptive statistics reflect the major decisions made and how SEM has been used in social work. Data were reported in aggregate, except for studies considered to be exemplary in a particular area.

## Results

### Overview of SEM studies reviewed

Table 2 shows a summary of journals by type of SEM. Eight of the ten journals under review published SEM research by social workers since 2001. The SEM studies in these journals included twenty confirmatory factor analyses, ten full SEM models, one path analysis and one latent growth curve model. *Research on Social Work Practice (RSWP)* published the most SEM studies in social work. All SEM articles in *RSWP* were confirmatory factor models.

### Study characteristics

Table 3 provides a summary of the study characteristics. AMOS (part of SPSS) was the most frequently used software, followed by LISREL, EQS and Mplus. Seven studies (21.9 per cent) failed to mention the software used. Twelve studies (37.5 per cent) employed a multi-group analysis to test for model invariance.

**Table 2** Number of published studies using structural equation modelling by journal title and model type

Journal	Path analysis	Confirmatory factor analysis	Full SEM	Latent growth curve
<i>Social Service Review</i>	–	–	2	–
<i>Social Work Research</i>	1	3	2	–
<i>Journal of Social Service Research</i>	–	2	2	–
<i>Child Abuse and Neglect</i>	–	–	1	–
<i>Research on Social Work Practice</i>	–	15	–	–
<i>Journal of Social Policy</i>	–	–	–	–
<i>American Journal of Community Psychology</i>	–	–	2	–
<i>Child and Youth Service Review</i>	–	–	–	1
<i>Families in Society</i>	–	–	–	–
<i>Journal of Community Psychology</i>	–	–	1	–
Total	1 (3.1%)	20 (62.5%)	10 (31.3%)	1 (3.1%)



**Table 3** Characteristics of studies using SEM

Study characteristics		N (%)
Software	AMOS	13 (40.6)
	LISREL	8 (25.0)
	EQS	3 (9.4)
	Mplus	1 (3.1)
	Not reported	7 (21.9)
Method of estimation	Maximum likelihood	21 (65.6)
	Unweighted least squares	3 (9.4)
	Not reported	8 (25.0)
Multiple-group method		12 (37.5)
Used dichotomous or ordinal measures	Yes	11 (34.3)
	No	20 (62.5)
Reported model modification	Indeterminable	1 (3.1)
	Yes	13 (40.6)
	No	19 (59.4)
Assessment of fit <sup>1</sup>	Chi-square only	0 (0)
	Goodness-of-fit indices only	12 (37.5)
	Chi-square and goodness-of-fit indices	19 (59.4)
	No measures of fit reported	1 (3.1)

<sup>1</sup> Ratio between chi-square and degrees of freedom was considered a goodness-of-fit index.

## Best practices

### Theoretical specification

One of the most important parts of SEM research—and arguably the most difficult—is the theoretical specification of a model. The accuracy of parameter estimates is partly dependent on the soundness of the theory and partly dependent on the validity of the measurement. Ten of the twelve studies (83.3 per cent) used path analysis or full SEM. All studies using path analysis or full SEM that included structural relations among observed or latent variables had a model with clear theoretical specifications.

### Competing models

Since all SEM models have equivalent models that generate the same predicted correlations or covariances (Kline, 2005), consideration of alternative models, especially equivalent models, can help researchers to rethink the substantive meaning of the favoured model. However, only five studies (15.6 per cent) in the sample used a competing models approach, which included testing a first-order factor model against higher orders, examining different numbers of first-order factors and looking at different structural formulation of models.

## Method of estimation

SEM uses fitting functions to minimize the difference between the population and sample. Fitting functions are a recipe to transform data into an estimate. In choosing a particular estimator, the main choice is between maximum likelihood and weighted least squares. Maximum likelihood, which assumes multivariate normality, is the default estimator in most SEM programmes.

Twenty-one (65.6 per cent) studies in the sample reported using maximum likelihood, three (9.4 per cent) reported unweighted least squares and eight (25.0 per cent) did not report a method of estimation. Twelve studies reported models that were based on dichotomous or ordinal data. Three studies that used unweighted least squares contained models with dichotomous or ordinal level variables; this is considered a proper method of estimation. Four studies that used maximum likelihood included dichotomous or ordinal data, which is considered an incorrect estimation method. Five studies that used dichotomous or ordinal data failed to report the method of estimation. Of these five studies, three used sample sizes that were less than 500, precluding the asymptotic distribution-free methods of estimation. All twenty-one studies using maximum likelihood and all eight studies that failed to indicate the method of estimation did not report an assessment of multivariate normality. It should be noted that multivariate normality is not assumed with unweighted least squares.

## Sample size

Sample size is an important consideration in every SEM analysis. There are many factors that need to be considered, which makes rules of thumb too general (see Muthén and Muthén, 2002). When data are non-normal or ordinal, different types of correlations (e.g. polychoric) and estimation methods (e.g. weighted least squares) are required. These situations typically require samples in excess of 1,000 cases, even with models having only a few free parameters (MacCallum *et al.*, 1996; Diamantopoulos and Siguaw, 2000).

In the studies under review, sample sizes ranged from 120 to 6,424 (mean = 850.22, median = 412.5, SD = 1,228.83). Only one study had a sample size that did not exceed 150 cases. Thirty-one studies included latent variables, twenty-nine of which reported the specific number of indicators per latent variable. Among these studies, the ratio of indicators per latent variable ranged from 2:1 to 28.3:1 (mean = 7.85, median = 5.75, SD = 7.02). Eleven studies (34.4 per cent) with latent variables had measurement models with as few as two or three indicators per latent variable. Due to identification requirements, the fit of these latent variable models could not be assessed independently of other models.

## Model fit

A number of goodness-of-fit indices have been developed to help researchers assess SEM models. However, none of these indices, except chi-square, has an associated significance test. Therefore, the assessment of model fit in SEM is more complicated than it is in other statistical approaches such as multiple regression (Shumacker and Lomax, 1996). Commonly used model fit indices include chi-square ( $\chi^2$ ), goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), root-mean-square residual (RMR) and the root mean square error of approximation (RMSEA). In addition, normed fit index (NFI) and comparative fit index (CFI) are created to allow a model comparison between a null model and a proposed model; normed chi-square (NC) and parsimonious fit index (PFI) are developed to assess model parsimony. Results of this study show that nineteen studies (59.4 per cent) under review reported both chi-square and multiple goodness-of-fit indices. Twelve studies reported goodness-of-fit indices only, and one study failed to report either chi-square or goodness-of-fit indices. All studies reviewed reported positive findings—that is, the theoretical model tested exhibited good fit with the data.

## Model modifications

It is common in SEM research to modify or re-specify models following results showing a poor statistical fit. It is very important that any changes to the model be theoretically based. Thirteen studies (40.6 per cent) reported modifying the original model. Five studies indicated adding correlated error terms as part of the modification process. For six studies, it was indeterminable at what stage of the process correlated error terms were added, but an assumption was made during the review process that these were part of model modifications, given the absence of any theoretical justification. One study did not report the number of correlated errors introduced. The number of correlated errors added as part of model modifications ranged from one to sixteen (mean = 6.4, median = 3.5, SD = 6.11). Four of the eleven studies that introduced correlated errors as part of the modification process provided a substantive justification—that is, due to item similarity. No other theoretical justification for adding correlated errors was observed. Thirteen studies (40.6 per cent) reported dropping items. All but one study reported the specific number of items dropped. The number of items dropped ranged from one to sixty-one (mean = 20.41, median = 10.5, SD = 22.5). The process for dropping items was guided by an exploratory factor analysis or confirmatory factor analysis (or both), using a particular threshold for factor loadings (e.g. < 0.60). Only one study provided a substantive justification for dropping items, which was attributed to an issue involving language translation.

## Reporting

According to Boomsma (2000) and Kline (2005), it is necessary to report sufficient information to allow evaluation and replication. This involves justifying SEM as appropriate for the research question under study, a path diagram to indicate hypothesized relationships between variables, the sampling procedure and the sample size, the choice of estimation procedure, a correlation/covariance matrix, overall model fit statistics, model parameter estimates and standard errors of parameter estimates that need to be reported because the underlying assumptions for each estimation method are different (Hoyle and Panter, 1995). Unstandardized coefficients need to be reported when making multiple group comparisons because standard deviations may differ across the groups (Kline, 2005).

In addition to the foregoing summary of gaps in reporting, only four studies (12.5 per cent) in the sample included a matrix that allowed for replication. Eight studies (25 per cent) did not report path or factor loadings, twenty-five studies (78.1 per cent) did not report error variances and eight studies (25 per cent) failed to report both.

## Discussion

This study reviewed how SEM is used by social work researchers, assessing the extent to which the use of SEM in social work research over a recent six-year period conformed to best practices. Results from the review show that a primary use of SEM in social work is employing confirmatory factor analysis to assess the psychometric properties of instruments. The journal *Research on Social Work Practice* was the primary outlet, publishing approximately two-thirds of the SEM research. Only a small proportion of the SEM research under review involved testing structural relations in a full SEM or path analysis or testing latent growth curves.

There were many important strengths and weaknesses in the studies reviewed. The authors of this study provide references to specific studies that are considered exemplars and have considerable promise for helping improve SEM in social work. References to studies that exhibited various problems have been intentionally excluded, as this study aims to raise awareness of various problems rather than criticize the work of others. A notable strength of the research was the empirical examination of some models and measures that have never been empirically tested and the generation of new insights into old topics through the use of SEM. Some studies took advantage of the flexibility of SEM by using a multiple group approach for testing model invariance across groups. A simple way to do this is to run a separate analysis for each group and compare unstandardized coefficients, while a more sophisticated way is to examine results of the model with constraints (Kline, 2005). In our sample, those studies involving

multiple groups appropriately used one of the two methods for group comparisons. The use of multiple groups can help specify the theory under testing. For instance, in the study by Mowbray *et al.* (2005), the model is examined and compared for African Americans and White Americans, which lends evidence to the cultural explanation of differential effects of race regarding how living arrangements affect mothers' well-being.

There was also a small set of studies that proposed and tested competing models, which are fundamental to theory building. The work of Benda (2002) was regarded as an exemplar. Specifically, Benda (2002) clearly described, tested and compared three different theoretical models of delinquency. The results had clear implications for theory and social work practice. The results could not have been nearly as rich if only a single model was tested. Two studies were also considered exemplars with respect to using SEM procedures that are unique to social work research methods. In particular, Bybee and Sullivan (2002) used a cross-lagged SEM model and Schmitz (2005) tested a latent growth curve. These studies can serve as models for working with longitudinal data.

There were problems in the SEM research that must be recognized to improve its quality. Foremost, many studies made significant modifications to the model with the expressed or implied purpose to improve model fit. This included dropping indicators, allowing cross-loadings and including numerous correlated error terms, which not only challenged the reliability and validity of some studies, but have made it impossible to replicate the findings with new data. It is well established in the SEM literature that empirically driven model modifications are likely to capitalize on chance variation in the data. Some studies explicitly stated or implied that the modification indices produced by the software were changes consistent with the theory, which is a serious misconception about the use of modification indices.

Studies that made significant modifications to the model also created barriers to meaningful interpretations, which were not observed in the reports. Specifically, how should either a researcher or practitioner interpret a model with numerous correlated error terms or cross-loadings or both? While these modifications have improved the empirical fit of the models, the consistency with theory must be discussed. Social work researchers employing SEM should use modification indices to inform potential problems with the theory, data and model. Any modifications made necessarily render the analysis exploratory instead of validating the theoretical model.

With the exception of the studies that tested competing models, there were no studies that reported final models that lacked a good fit with the data. This may be one of the reasons driving the extensive number of model modifications to improve model fit. It is important to recognize negative findings if they are important. On the other hand, researchers employing SEM should recognize the knowledge gains of both positive and negative findings prior to carrying out any SEM analysis. This position

can also help research demarcate where hypothesis testing has concluded and exploratory analysis has begun.

Another issue of concern is the lack of attention that is paid to the method of estimation. The majority of the studies employed the default estimation method (i.e. maximum likelihood) or failed to indicate the method of estimation. There was almost no attention paid to the assumptions of estimation, suggesting that many researchers have overlooked this part of the analysis. Researchers using maximum likelihood with non-normal and non-continuous data can have severely biased results. In some cases, if a more suitable method of estimation were used, the efforts paid to model modifications may have been unnecessary.

Another problem with the studies reviewed was the heterogeneity and problems with reporting, which also created challenges to the review process. In particular, the process by which modifications were made was unclear. This was especially common in studies that added correlated errors to improve model fit. It was not clear if all the correlated error terms were at one time based on the full list of modifications reported by the software, or if one modification was made, the model tested and then another one added. This is an important consideration, as one approach can lead to different results from the other approach.

Although SEM is clearly recognized as an important analytic tool in social work research, it is still not being used to its fullest potential. For example, during the retrieval process of this study, a number of studies were observed that could have been stronger by using SEM as opposed to other analytic techniques. Instead of using SEM, some recent studies used ordinary least squares regression to conduct a path analysis, which in fact can be more appropriately done by SEM because it can reduce type I error by testing all the structural relations simultaneously. Some studies have also examined the factor structure of existing measures using traditional exploratory factor analysis without taking into account measurement error. Since these measures are already developed and in use, a more suitable technique would be a test of the factor structure using SEM.

Social work researchers using SEM should also consider many of its options that are not being used to their fullest extent. For example, social work theories often imply reciprocal or feedback relationships. Although close attention must be paid to issues of model identification, these can be specified and tested as non-recursive relationships, which were not present in any of the studies reviewed.

## Limitations

The foregoing results and discussion should be considered within the study limitations. This study reviewed a set of top social work journals. Although using the reputation approach overcomes problems associated with the

validity of impact indicators, it did exclude other journals that are of high quality and regarded as important publication outlets among social work researchers that are involved in inter-disciplinary initiatives. Although one British journal was included in this review (i.e. the *Journal of Social Policy*), the remaining journals were US publications and may not reflect how SEM is used among the broader community of social work researchers. At the same time, all the journals reviewed have an international readership that is influenced by research published. As the best practices of SEM are universal, the results are applicable to the international community but may not be generalizable beyond the journals reviewed.

## Future research

The current study provided an assessment of data preparation, including sample size, random sample, multivariate normality and multivariate outcome. A more comprehensive assessment should also give critical attention to how missing data are handled. This study did not assess for proper theoretical specification of latent variables—in particular, making a proper distinction between formative and reflective indicators (see Bollen and Lennox, 1991). Numerous studies appeared to exhibit specification problems, which suggest further critical assessment is necessary in judging the quality of SEM social work research. In addition, although SEM is frequently referred to as ‘causal modelling’, without experimental design or strict assumptions, it is hard to know whether fitting a structural equation model actually tells anything valid about causal effects (Holland, 1988; Pearl, 1998). While the issue of causality is complicated and beyond the scope of this study, it can be further examined in future research.

As noted earlier, the studies reviewed showed significant heterogeneity and gaps in reporting. Although the reliability of the coding system was checked, there were a number of judgment calls that had to be made, as the methodology was unclear. Thus, the lack of clarity in reporting may also have influenced the reliability of the review process. For example, some studies employing full SEM reported an initial test of the measurement models, but it was not clear whether item parcels or averaged scores were used in the final analysis. To help overcome this problem, the authors used a group discussion to address uncertainties in the coding process.

## Conclusions

Valid and reliable research results depend on a method appropriate to the research problem and on that method being used correctly. The results of this study identified exemplars of SEM practices in social work.

However, there were many instances of poor SEM practices, which compromise the social work knowledge base. In order for SEM to benefit social work researchers maximally, it is imperative that best practices are followed. Anything less hampers the development of knowledge, which, in turn, undermines effective policy and service delivery. A number of excellent introductory SEM textbooks, the SEM listserv and SEMNET serve as rich resources for becoming informed about SEM best practices, learning new advances and seeking technical assistance.

## References

- Benda, B. B. (2002) 'A test of three competing theoretical models of delinquency using structural equation modeling', *Journal of Social Service Research*, **29**(2), pp. 55–91.
- Bollen, K. A. and Lennox, R. (1991) 'Conventional wisdom on measurement: A structural equation perspective', *Psychological Bulletin*, **110**(2), pp. 305–14.
- Boomsma, A. (2000) 'Reporting analyses of covariance structures', *Structural Equation Modeling*, **7**, pp. 461–83.
- Breckler, S. J. (1990) 'Application of covariance structure modeling in psychology: Cause for concern?', *Psychological Bulletin*, **107**(2), pp. 260–372.
- Bybee, D. I. and Sullivan, C. M. (2002). 'The process through which an advocacy intervention resulted in positive change for battered women over time', *American Journal of Community Psychology*, **30**(1), pp. 103–32.
- Diamantopoulos, A. and Siguaw, J. A. (2000) *Introducing LISREL: A Guide for the Uninitiated*, London, Sage Publications, Inc.
- DiStefano, C. and Hess, B. (2005) 'Using confirmatory factor analysis for construct validation: An empirical review', *Journal of Psychoeducational Assessment*, **23**, pp. 225–41.
- Gillespie, D. F., Alsup, R. and Rubio, D. M. (1995) 'Structural equation modeling: An innovation for doctoral social work curricula', *Journal of Social Work Education*, **31**, pp. 219–28.
- Hamilton, M. A. and Orme, J. G. (1990) 'Examining the construct validity of three measures of parenting knowledge using LISREL', *Social Service Review*, **14**, pp. 1–20.
- Holland, P. W. (1988) 'Causal inference, path analysis, and recursive structural equations models (with discussion)', in Clogg, C. C. (ed.), *Sociological Methodology 1988*, Washington, DC, American Sociological Association, pp. 449–93.
- Hoyle, R. H. and Panter, A. T. (1995) 'Writing about structural equation models', in Hoyle, R. H. (ed.), *Structural Equation Modeling: Concepts, Issues, and Applications*, Thousand Oaks, CA, Sage Publications, pp. 158–76.
- Jenson, J. M. (2005) 'Structural factors and the quality of publication in social work journals', *Social Work Research*, **29**(2), pp. 67–71.
- Kline, R. B. (2005) *Principles and Practice of Structural Equation Modeling*, 2nd edn, New York, Guilford Press.
- MacCallum, R. C. and Austin, J. T. (2000) 'Applications of structural equation modeling in psychological research', *Annual Review of Psychology*, **51**, pp. 201–26.



- MacCallum, R. C., Browne, M. W. and Sugawara, H. M. (1996) 'Power analysis and determination of sample size for covariance structure modeling', *Psychological Methods*, **1**, pp. 130–49.
- Martens, M. P. (2005) 'The use of structural equation modeling in counseling psychology research', *The Counseling Psychologist*, **33**(3), pp. 269–98.
- Mowbray, C. T., Bybee, D., Hollingsworth, L., Goodkind, S. and Oyserman, D. (2005) 'Living arrangements and social support: Effects on the well-being of mothers with mental illness', *Social Work Research*, **29**(1), pp. 41–55.
- Muthén, L. K. and Muthén, B. O. (2002) 'How to use a Monte Carlo study to decide on sample size and determine power', *Structural Equation Modeling*, **4**, pp. 599–620.
- Orme, J. G. (1993) 'Structural equation modeling: An introduction to LISREL' (Invited Faculty Development Institute), Council on Social Work Education Annual Programme Meeting, New York, NY.
- Orme, J. G. (1995) 'Structural equation modeling with LISREL and SIMPLIS' (Invited Workshop), Society for Social Work and Research Annual Conference, Washington, DC.
- Orme, J. G. and Fickling, J. A. (1992) 'A preliminary examination of the Knowledge of Normative Infant Development (KNID) Inventory using structural equation modeling', *Journal of Social Service Research [Special Issue on Quantitative Methods in Social Work Research]*, **17**, pp. 59–86.
- Pearl, J. (1998) 'Graphs, causality, and structural equation models', *Sociological Methods and Research*, **27**, pp. 226–84.
- Russell, D. W. (2002) 'In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*', *Personality and Social Psychology Bulletin*, **28**(12), pp. 1629–46.
- Schmitz, M. F. (2005) 'Effects of childhood foster care and adoption on adulthood child-bearing', *Child and Youth Services Review*, **27**, pp. 85–98.
- Schumacker, R. E. and Lomax, R. G. (1996) *A Beginner's Guide to Structural Equation Modeling*, Lawrence Erlbaum Associates, Mahwah, NJ.
- Sellers, S. L., Mathiesen, S., Perry, R. and Smith, T. (2004) 'Evaluation of social work journal quality: Citation vs reputation approaches', *Journal of Social Work Education*, **40**(1), pp. 143–60.
- Sellers, S. L., Mathiesen, S. G., Smith, T. and Perry, R. (2006) 'Perceptions of professional social work journals: Findings from a national survey', *Journal of Social Work Education*, **42**(1), pp. 139–60.
- Society for Social Work Research (SSWR) (2005) 'Peer review and publication standards in social work journals: The Miami statement', *Social Work Research*, **29**(2), pp. 119–121.
- Tomarken, A. J. and Waller, N. G. (2005) 'Structural equation modeling: Strengths, limitations, and misconceptions', *Annual Review of Clinical Psychology*, **1**, pp. 31–65.

## Appendix: Studies included in systematic review

- Abbott, A. A. (2003) 'A confirmatory factor analysis of the Professional Opinion Scale: A values assessment instrument', *Research on Social Work Practice*, **13**(5), pp. 641–66.
- Abell, N., Ryan, S. and Kamata, A. (2006) 'Assessing capacity for self-care among HIV positive heads of household: Bilingual validation of the parental self-care scale', *Social Work Research*, **30**(4), pp. 233–43.
- Abell, N., Ryan, S., Kamata, A. and Citrolo, J. (2006) 'Bilingual validation of the family responsibility scale: Assessing stress among HIV+ heads of household', *Journal of Social Service Research*, **32**(2), pp. 195–212.
- Begun, A. L., Murphy, C., Bolt, D., Weinstein, B., Short, L. and Shelley, G. (2003) 'Characteristics of the Safe At Home Instrument for Assessing Readiness to Change Intimate Partner Violence', *Research on Social Work Practice*, **13**, pp. 65–79.
- Benbenishty, R., Astor, R. A., Zeira, A. and Vinokur, A. D. (2002) 'Perceptions of violence and fear of school attendance among junior high school students in Israel', *Social Work Research*, **26**(2), pp. 71–87.
- Benda, B. B. (2002a) 'Test of a structural equation model of comorbidity among homeless and domiciled military veterans', *Journal of Social Service Research*, **29**(1), pp. 1–35.
- Benda, B. B. (2002b) 'A test of three competing theoretical models of delinquency using structural equation modeling', *Journal of Social Service Research*, **29**(2), pp. 55–91.
- Bennett, M. D. and Miller, D. B. (2006) 'An exploratory study of the Urban Hassles Index: A contextually relevant measure of chronic multidimensional urban stressors', *Research on Social Work Practice*, **16**(3), pp. 305–14.
- Bybee, D. I. and Sullivan, C. M. (2002) 'The process through which an advocacy intervention resulted in positive change for battered women over time', *American Journal of Community Psychology*, **30**(1), pp. 103–32.
- Choi, Y., Golder, S., Gillmore, M. R. and Morrison, D. M. (2005) 'Analysis with missing data in social work research', *Journal of Social Service Research*, **31**(3), pp. 23–48.
- Dia, D., Smith, C. A., Cohen-Callow, A. and Bliss, D. L. (2005) 'The Education Participation Scale-Modified: Evaluating a measure of continuing education', *Research on Social Work Practice*, **15**(3), pp. 213–22.
- Early, T. J., Gregoire, T. K. and McDonald, T. P. (2001) 'An assessment of the utility of the Child Behavior Checklist/4–18 for social work practice', *Research on Social Work Practice*, **11**, pp. 597–612.
- Kahng, S. K. and Mowbray, C. (2004) 'Factors influencing self-esteem among individuals with severe mental illness: Implications for social work', *Social Work Research*, **28**(4), pp. 225–36.
- Landsman, M. J. (2001) 'Commitment in public child welfare', *Social Service Review*, **75**(3), pp. 386–419.
- Lau, A. and Takeuchi, D. (2001) 'Cultural values in helpseeking for child behavior problems: Value orientation, affective responding, and severity appraisals among Chinese American parents', *Journal of Community Psychology*, **29**, pp. 675–92.
- Macgowan, M. J. and Newman, F. L. (2005) 'Factor structure of the group engagement measure', *Social Work Research*, **29**(2), pp. 107–18.
- McMillan, J. D. and Abell, N. (2005) 'Validating the Level of Stability Index for Children', *Research on Social Work Practice*, **16**(3), pp. 326–37.

- Mowbray, C. T., Bybee, D., Hollingsworth, L., Goodkind, S. and Oyserman, D. (2005) 'Living arrangements and social support: Effects on the well-being of mothers with mental illness', *Social Work Research*, **29**(1), pp. 41–55.
- Nugent, W. R. (2004) 'A validity study of scores from self-anchored-type scales for measuring depression and self-esteem', *Research on Social Work Practice*, **14**(3), pp. 171–9.
- O'Hare, T., Tran, T. V. and Collins, P. (2002) 'Validating the Internal Structure of the Practice Skills Inventory', *Research on Social Work Practice*, **12**, pp. 653–68.
- O'Keefe, M., Mennen, F. and Lane, C. J. (2005) 'An examination of the factor structure for the youth self report on a multiethnic population', *Research on Social Work Practice*, **16**(3), pp. 315–25.
- Schmitz, M. F. (2005) 'Effects of childhood foster care and adoption on adulthood child-bearing', *Child and Youth Service Review*, **27**, pp. 85–98.
- Sutherland, C. A., Bybee, D. I. and Sullivan, C. M. (2002) 'Beyond bruises and broken bones: The joint effects of stress and injuries on battered women's health', *American Journal of Community Psychology*, **30**(5), pp. 609–36.
- Tran, T. V., Ngo, D. and Conway, K. (2003) 'A cross-cultural measure of depressive symptoms among Vietnamese Americans', *Social Work Research*, **27**(1), pp. 56–65.
- Tyson, E. H. (2006) 'Rap-music Attitude and Perception Scale: A validation study', *Research on Social Work Practice*, **16**(2), pp. 211–23.
- Wilke, D. J., Kamatab, A. and Cash, S. J. (2005) 'Modeling treatment motivation in substance-abusing women with children', *Child Abuse and Neglect*, **29**, pp. 1313–23.
- Yoo, J. and Brooks, D. (2005) 'The role of organizational variables in predicting service effectiveness: An analysis of a multilevel model', *Research on Social Work Practice*, **15**(4), pp. 267–77.