

Mediational Analysis in HIV/AIDS Research: Estimating Multivariate Path Analytic Models in a Structural Equation Modeling Framework

Angela Bryan · Sarah J. Schmiede ·
Michelle R. Broaddus

Published online: 18 August 2006
© Springer Science+Business Media Inc. 2006

Abstract Mediational analyses have been recognized as useful in answering two broad questions that arise in HIV/AIDS research, those of theoretical model testing and of the effectiveness of multicomponent interventions. This article serves as a primer for those wishing to use mediation techniques in their own research, with a specific focus on mediation applied in the context of path analysis within a structural equation modeling (SEM) framework. Mediational analyses and the SEM framework are reviewed at a general level, followed by a discussion of the techniques as applied to complex research designs, such as models with multiple mediators, multilevel or longitudinal data, categorical outcomes, and problematic data (e.g., missing data, nonnormally distributed variables). Issues of statistical power and of testing the significance of the mediated effect are also discussed. Concrete examples that include computer syntax and output are provided to demonstrate the application of these techniques to testing a theoretical model and to the evaluation of a multicomponent intervention.

Keywords Mediation · Path analysis · Structural equation modeling (SEM) · Multicomponent interventions

Introduction

The advent of accessible structural equation modeling (SEM) programs (e.g., AMOS, MPlus, EQS, and LISREL) in combination with the focus on theory testing and the mechanisms of behavior change of the HIV/AIDS field, has caused an explosion in the use of SEM to test theory-based mediational questions. The result has been that researchers who may never have had training in mediational analysis or SEM per se are not only being asked to read and understand such analyses but are increasingly being asked to *produce* these analyses themselves. The goal of this paper, therefore, is to serve as a primer of the use of path analysis within the framework of SEM to test mediational questions that may arise in HIV/AIDS research. We note at the outset that our use of the term SEM denotes a class of analytic techniques that usually include the estimation of unobserved or *latent* constructs and an estimation of the structure of the relationships among these latent constructs (Loehlin, 1992). We restrict this paper to the special case of SEM known as path analysis (Marcoulides & Schumacker, 1996) wherein every variable in the model is directly measured or *observed*. We do this for two reasons. First, the mediational tests we describe become much more complex with the addition of latent variables (see Kenny, 2006). Second, in our experience, it is often difficult to fit complex multicomponent latent variable models with the sample sizes common in most applied research applications, and thus the much more common approach is to fit a path model with observed variables. Thus, this is the approach we recommend when the number of components is large. An important caveat to this recommendation, however, is that

A. Bryan (✉) · S. J. Schmiede · M. R. Broaddus
Department of Psychology, University of Colorado,
Muenzinger Psychology Building, UCB 345, Boulder, CO
80309-0345, USA
e-mail: angela.bryan@colorado.edu

researchers should take care that they are using the most reliable measures possible when estimating such models, as a key disadvantage of dropping from a latent variable to a measured variable approach is that measured variable approaches cannot correct for unreliability of measurement.¹

Some who are familiar with traditional mediational analyses via OLS or logistic regression might wonder why we approach this question from an SEM framework when it is possible to piece together the findings from a number of regression runs to learn essentially the same information. We agree that in the case that a researcher has one independent variable, one dependent variable, one mediator, very little or no missing data, and a normal distribution on all three variables, the use of multiple regression runs and path analysis is identical. However, there are a number of reasons that we recommend an SEM framework over OLS or logistic regression for the types of applications common in applied HIV/AIDS research. Path models allow one to examine direct, indirect, and total effects simultaneously in one model, they allow the testing of multiple mediators/dependent variables and complex mediational chains, and they allow the testing of specific indirect effects within those complex chains. Additionally, it is easier to apply some of the bootstrap resampling techniques in SEM programs, and such programs are also able to appropriately correct for missing data problems and non-normality in the data. Finally, from a theoretical perspective, it is more satisfying to simultaneously test the validity of the entire theory as hypothesized in the context of a structural model.

It is our hope that this paper, while not an exhaustive treatment of the use of path analysis via SEM for testing questions of mediation, will serve as a baseline level of knowledge for those wishing to interpret and understand such models and as a “launch pad” of sorts for those wishing to use these techniques in their own research.

Examples from Recent HIV/AIDS Empirical Work

In order to demonstrate common situations in which the SEM framework may be useful for answering questions related to mediation, we begin with representative examples from the literature of two different

forms: one in which path analysis was used to test mediational questions in a purely theoretical context (Wayment et al., 2003) and one in which tests for the mechanisms of change via an intervention were accomplished with path analysis (Bryan, Aiken, & West, 1996; c.f. West & Aiken, 1997; Kraemer, Wilson, Fairburn, & Agras, 2002).

Wayment et al. (2003) utilized path analysis in order to test the extent to which the effects of sociocultural variables on risky sexual behavior were mediated by Health Belief Model (Rosenstock, 1990) variables. In other words, they hypothesized that variables such as marital status influence one’s health beliefs, and that health beliefs, in turn, influence one’s behavior. Data were collected from a randomly selected sample of 260 single versus married or cohabiting White women in face-to-face interviews. Measures included demographic information (e.g., women’s age, marital status, SES, and religiosity), as well as Health Belief Model constructs of perceived susceptibility or invulnerability to HIV and STD, perceived seriousness of STD infection, AIDS transmission and STD prevention knowledge, barriers to STD checkups, perceived condom efficacy against pregnancy and STDs, efficacy of the contraceptive pill to prevent pregnancy, and health locus of control. The outcomes of interest included a risky sex index (participation in sexual behaviors including vaginal/anal intercourse, sex without a condom, fellatio, etc.), number of unintended pregnancies, number of sexual partners, and method of birth control used (barrier method [i.e., condoms] versus nonbarrier or no method).

The authors correctly point out that due to the cross-sectional nature of the study they are presenting a *possible* causal sequence based on plausibility and previous theoretical research. This is thus a model of postdiction (c.f., Albarracin, Fishbein, & Middlestadt, 1998)—a phenomenon we explore in more depth later in the paper. They present a final model that provides a good fit to the data, and their analyses indicate partial support for their hypotheses. For example, higher religiosity was associated with decreased beliefs in pill efficacy for pregnancy prevention, which was in turn related to higher numbers of unintended pregnancies. The effect of marital status on outcomes was partially mediated by perceived susceptibility to HIV, such that married or cohabiting women felt lower perceived susceptibility and, in turn, had fewer sex partners. Some hypotheses were completely unsupported—for example, there were no significant predictors of the risky sex index, and health locus of control and knowledge of STD prevention did not relate to any predictor variable or outcome.

¹ A key advantage of using latent variables instead of measured variables is that latent variables account for unreliability of measurement (Jaccard & Wan, 1995). When measures are unreliable, the regression coefficients/path coefficients are biased. For more information on tests of mediation within a latent variable framework, see Hoyle and Kenny (1999).

While Wayment et al.'s (2003) focus was the testing of a theoretical model, Bryan et al. (1996) used path analysis to examine the effects of a condom promotion intervention on condom use intentions in a sample of college women, as mediated by theory-based model variables that the intervention was designed to change. The authors developed a theoretical model of condom use specifically tailored to young women. It was hypothesized that the intervention (versus control) would change perceptions of sexuality for women (acceptance of sexuality and control over the sexual encounter), perceived susceptibility to and severity of common STDs (i.e., chlamydia and gonorrhea), and perceived benefits of and self-efficacy for condom use (obtaining, negotiating with a partner, dealing with subsequent partner dissatisfaction, and mechanics of use).

Participants completed four assessments: baseline, post-intervention, 6 week follow-up, and 6 month follow-up. Baseline and post-intervention assessments measured all model constructs. Follow-up assessments measured behavior in the period since last assessment. In a MANCOVA of intervention effects on model constructs, all variables except susceptibility and severity were significantly impacted by the intervention and were therefore maintained in the structural equation modeling analyses. The final model indicated that the intervention had a direct impact on self-efficacy over and above the mediated paths through control and acceptance, and that the intervention had an indirect effect on attitudes that was mediated by its effects on perceived benefits of condoms. Attitudes and self-efficacy, in turn, affected intentions to use condoms. Although it was not possible to analyze a model of condom use *behavior* over time due to the low incidence of sexual intercourse within the six month period of the study, using path analytic techniques within an SEM framework allowed the authors to show that a theoretically-based intervention affected intentions to use condoms through the hypothesized mechanisms. The analyses in this paper also demonstrated theoretical model constructs that were *unaffected* by the intervention; i.e., perceived susceptibility to and severity of common STDs.

The Wayment et al. (2003) and Bryan et al. (1996) studies demonstrate the two most common applications of mediational analysis via path analysis in HIV/AIDS research: theory-testing and probing the effectiveness of individual intervention components. These two examples can be further distinguished in terms of whether or not participants were randomly assigned to condition: models that are used to test a theory are typically based on non-experimental research, as

compared to models in which participants are randomly assigned to experimental condition in order to determine mechanism. We next discuss issues related to inferring causality in these models, particularly in instances where there is not random assignment.

Causality and the SEM Framework

The most serious mistake in SEM analyses, or any other analytic technique for that matter, is to infer causality when one has no grounds to do so (c.f., Pearl, 2000). Evidence for causality is certainly strongest in experimental designs with random assignment, although even experimental designs do not provide absolute proof of causality. Many applications of SEM have neither random assignment nor experimental manipulation of treatment—any appeal to causality in such models is thus tenuous at best (Shadish, Cook, & Campbell, 2002). The connection between SEM and mistaken notions about causal inference may be because of the use of the term “causal model” in the historical SEM literature (c.f., Bentler, 1980). Any introductory text on SEM (e.g., Bollen, 1989; Loehlin, 1992; an edited volume by Hoyle, 1995) will caution in the first chapter that the term “causal model” is a misnomer for this class of analytic techniques. Analyzing one's data in an SEM framework based on theories of causation, can (under the right circumstances) *imply* causation, and can be used to generate causal hypotheses. But analyzing mediational questions via SEM in a cross-sectional, observational data set can never provide strong evidence of causation, nor should such language be used in the discussion of such models.

One circumstance under which weak claims for causal inference can be made in the context of non-experimental data is when there is temporal precedence between the predictor and criterion variables. This is particularly relevant to mediational designs because the mediator and outcome variables are not experimentally manipulated, even in situations of random assignment to levels of the independent variable. Thus, even though it may be reasonable to claim evidence of causality for the relationship between the independent variable and the mediator and the independent variable and the outcome, causal claims regarding the test of the mediator on the outcome are severely limited by several alternative explanations (e.g., the outcome may instead lead to the mediator; a third variable may explain levels on both the mediator and the outcome). Evidence for causal direction between the mediator and the outcome will certainly be bolstered if the mediator precedes the outcome variable in time.

A related issue is that, when all variables are measured contemporaneously, there may be one or more plausible configurations of variables as an alternative to any given SEM model that may fit the data as well as or identically to the final, reported model. As an example, many researchers are very surprised that when they flip the mediator and the outcome, they obtain reasonable results. MacCallum, Wegener, Uchino, and Fabrigar (1993) define equivalent models as those that exhibit identical fit to the data as the original model, but that have different patterns of relationships among variables compared to the original model. The authors conducted a review of the literature in which they examined the prevalence of equivalent models in 53 published studies, finding that alternative equivalent models existed for 90% of the studies. Plausible equivalent models may be particularly likely when the research design is cross-sectional and/or non-experimental. For example, in the Wayment et al. (2003) study we described, the effect of marital status on number of partners was found to be mediated through perceived susceptibility to HIV. Alternatively, it is plausible that marital status is the underlying reason for a decreased number of partners, and as a result, married participants have lower perceived susceptibility. Although this is indeed a problem in SEM research, MacCallum et al. (1993) note that a first step at a solution is for researchers to be cognizant of the fact that equivalent models likely exist for their model and to critically evaluate whether the alternative models might be justified substantively. Clear arguments for the chosen model should be provided. As a reader of the Wayment et al. (2003) paper would see, the authors in this case clearly state the limitations imposed on their findings by the cross-sectional design, and make no claims to causality, as is appropriate.

Another issue related to the mistaken inference of causality in tests of multicomponent models is the familiar strategy of HIV/AIDS researchers to present models in which behavior, measured contemporaneously with all other model constructs, serves as the outcome variable. So as not to pick on any group of researchers unnecessarily, we cite an example from our own work. In Bryan, Fisher, Fisher, and Murray (2000), we utilized the Information-Motivation-Behavioral skills model (Fisher & Fisher, 2002) to examine high risk sexual behavior among intravenous drug users in methadone treatment. In this model, our outcome was condom use behavior, and the predictors were information, motivation, and behavioral skills, though we measured all of these variables *at the same time*. Albarracin et al. (1998) review the problematic nature of such models of “postdiction” (as opposed to

prediction), in that they are actually using currently held cognitions and beliefs to account for variability in *past* behavior. So from a temporal precedence perspective, the behavior happened before the assessment of current information, motivation, skills, etc. and thus these variables simply cannot logically “predict” behavior that has already happened. These models often result in the overestimation of the ability of model constructs to account for variability in behavior.

Classic causal inference aside, various theories specify the role of past behavior in models of future behavior. Fishbein and Ajzen (1975) suggest that the effects of learning (i.e., past behavior) on action are expected to be indirect and to occur through specific changes in beliefs and evaluations. Thus, past behavior is an exogenous variable that “predicts” current cognitions and beliefs. Experimental work confirms these ideas. For example, Albarracin and Wyer (2000) showed that past behavior influenced future behavior through an indirect route mediated by attitude change. In general we recommend that if researchers place previous behavior as an outcome variable in a cross-sectional mediational model, they be extremely cautious in the interpretation of the findings with such models and note clearly in the limitations section of such work that these are postdiction models.

With these cautions regarding causality firmly in mind, we now turn to a brief history of mediational analysis and review of the current “state of the art” in the methodological and quantitative literature on mediation. Although the following section is somewhat technical in nature, the methodology will become clearer to the reader further in the paper when we use specific examples from our own research to demonstrate the application of the procedures. Note that in our general description of mediation, we will use the terms independent variable, mediator, and dependent variable. The use of these terms is not entirely consistent with non-experimental, cross-sectional data (e.g., Wayment et al., 2003), and technically predictor and outcome variable would be preferred under these circumstances. However, for simplicity, we will use independent and dependent variable to describe the first variable in a mediational chain and the outcome variable in a mediational chain, respectively, regardless of the design under consideration.

Background of Mediation

Historically, mediation has been conceptualized and measured in various ways across disciplines, with the most common approach in the psychological literature

stemming from the work of Kenny and colleagues (Baron & Kenny, 1986; Judd & Kenny, 1981). According to the framework set forth in these articles, a series of causal steps are required to conclude that mediation has occurred: (a) the independent variable must significantly impact the dependent variable (τ in Fig. 1a); (b) the independent variable must have a significant effect on the mediator (α in Fig. 1b); (c) the mediator must significantly impact the dependent variable (β in Fig. 1b), adjusting for scores on the independent variable; and, (d) the effect of the independent variable on the dependent variable should decrease when the mediator is included in the equation (τ' in Fig. 1b).

The contribution of this causal steps method to a theoretical understanding of mediation has been widely acknowledged (MacKinnon, 1994, 2000) and the causal steps method is still the most commonly used approach in social science journals (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Preacher & Hayes, 2004). Because the logic underlying the causal steps approach is critical for an understanding of mediation analyses, we first describe a basic, general mediational model. We then move to a discussion of several extensions beyond this method that are particularly relevant for mediation within the path analytic framework. Such extensions include: testing the significance of the mediated effect; statistical power; models including multiple mediators; and extensions beyond ordinary least squares (OLS) regression (e.g., multilevel models, longitudinal models, categorical outcomes).

Basic Mediational Model

A series of OLS regression equations are sufficient to answer the main questions about mediation in the case of simple models with one independent variable, one dependent variable, and one mediator, regardless of the method used for the determination of the presence of mediation:

$$Y = \beta_{0(1)} + \tau X + e_{(1)} \quad (1)$$

$$M = \beta_{0(2)} + \alpha X + e_{(2)} \quad (2)$$

$$Y = \beta_{0(3)} + \tau' X + \beta M + e_{(3)} \quad (3)$$

In these equations, X represents the independent variable, Y represents the dependent variable, and M represents the mediating variable (see Fig. 1a, b). The model intercepts for each of the three equations are represented by $\beta_{0(1)}$, $\beta_{0(2)}$, and $\beta_{0(3)}$, respectively, and the residuals by $e_{(1)}$, $e_{(2)}$, and $e_{(3)}$. Equation 1 shows the

direct effect (τ) of the independent variable (X) on the dependent variable (Y). Equation 2 represents the direct effect of X on M (α). Equation 3 shows the effect of X on Y (τ') when the effect of the mediator on Y (β) is included in the model. These equations follow from the logic of the causal steps method, such that Eq. 1 tests Step 1, Eq. 2 tests Step 2, and Eq. 3 tests Steps 3 and 4. Two alternate methods of assessing mediation—the difference of coefficients method and the product of coefficients method—extend the causal steps method in that the same equations can be used to test the size and significance of the mediated effect. The difference in coefficients method obtains an estimate of the mediated effect ($\tau - \tau'$) by using Eqs. 1 and 3 to measure the difference in the effect of X on Y with and without M included in the model. The product of coefficients method obtains an estimate ($\alpha\beta$) by using Eqs. 2 and 3 as a product of the effect of X on M (α) and the effect of M on Y (β). In most situations (e.g., continuous outcomes, no clustering in the data), the difference of coefficients and product of coefficients methods will provide identical estimates of the mediated effect (MacKinnon, Warsi, & Dwyer, 1995). The precise way in which these point estimates can be used to test the significance of the mediated effect will be addressed further in the paper.

The above equations can be related to the three broad types of effects that are examined in multi-component path analytic models (Bollen, 1987). Direct effects model the influence of one variable on another, unmediated by any other variables (represented by τ' from Eq. 3 above and depicted in Fig. 1b). Indirect effects represent influences of one variable on another that occur through one or more other variables (represented by $\alpha\beta$ from Eqs. 2 and 3 above and depicted in Fig. 1b). Total effects represent the sum of the direct and indirect effects (represented either by τ from Eq. 1 and depicted in Fig. 1a or by $\tau' + \alpha\beta$ from Eqs. 2 and 3 and depicted in Fig. 1b). These three types of effects deserve mention in order to demonstrate how the three

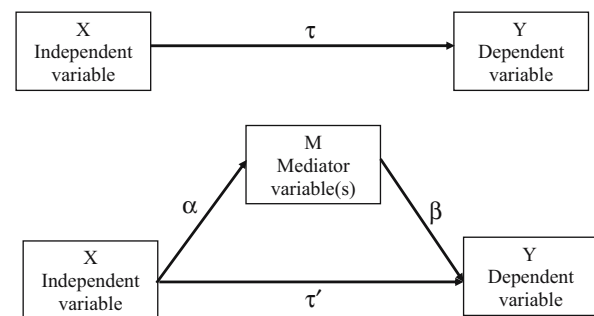


Fig. 1 Basic mediational model

broad methods of testing for mediation (causal steps, difference of coefficients, and product of coefficients) translate to an SEM context. The product of coefficients method is the most easily applied in an SEM framework because direct, total, and indirect effects can all be examined in a single model (Fig. 1b) that is based on the two equations of the product of coefficients method (Eqs. 2 and 3). In contrast, two separate models (Fig. 1a, b) and all three equations would be required to obtain direct, indirect, and total effect estimates if one of the other two methods (causal steps or the difference of coefficients) were applied to an SEM context. The main focus of this paper will thus be on the product of coefficients method (also termed the “indirect effects” method).

At this point, it is also necessary to clarify the terminology we will use, as mediation has been defined somewhat inconsistently in the literature. Some have reserved the term “mediation” for the causal steps approach, using the term “intervening variables” or “indirect effects” for methods in which the initial direct path from X to Y isn’t explicitly measured (i.e., the product of coefficients method) or is not significant (Holmbeck, 1997; Preacher & Hayes, 2004). However, others have noted that a significant direct path from the independent variable to the dependent variable may not be necessary, or even realistic, to imply mediation (Collins, Graham, & Flaherty, 1998; MacKinnon, 2000; Shrout & Bolger, 2002). For example, there tends to be low power for detecting the direct effect of X on Y (MacKinnon et al., 2002), and the requirement of a significant X to Y path rules out models with inconsistent indirect effects, for example, models where X positively impacts M , which, in turn, negatively impacts Y (MacKinnon, Krull, & Lockwood, 2000). For our purposes, we use the term “mediation” at a more general level to refer to any situation in which at least one variable is intermediate within a hypothesized causal chain, regardless of whether there is an initially significant direct effect of X on Y . Consequently, the terms “indirect effect” and “mediated effect” will be used interchangeably throughout the remainder of the paper.

Extensions/Advantages of Mediation Analysis

Significance of the Mediated Effect

Many in the field are intimately familiar with and generally practice the causal steps method. This series of tests is suggestive of mediation, but the use of these tests alone is limited because it does not explicitly

provide an estimate of the size of the mediated effect or its standard error, preventing examination of the significance of, and the confidence intervals around, the mediated effect. Preacher and Hayes (2004) note that using the causal steps method without formal tests of the mediated effect may lead to erroneous conclusions, and give examples of situations where reliance on the causal steps method alone may result in either Type I or Type II errors. MacKinnon (1994) also recommends formal significance tests of the mediated effect, noting that it is possible for the statistical test of the mediated effect to be nonsignificant, even when there is a significant effect of the independent variable on the mediator and a significant effect of the mediator on the outcome. Such a situation is problematic, as there may be several alternative interpretations of the results in addition to mediation (e.g., model misspecification, suppressor effects, lack of causal relationship between X and Y through the mediator). We thus believe that it is crucial to test whether or not the size of the mediated or indirect effect is significantly different from zero.

A common way of calculating the significance of the mediated effect is by dividing the mediated effect estimate (either $\tau - \tau'$ or $\alpha\beta$) by one of several standard error estimates, with the resulting coefficient distributed as a z -statistic (where absolute values greater than 1.96 indicate significance at an α of .05). Sobel (1982) provides a complete description of his derivation of the standard error of a mediated effect via the use of the multivariate delta method, such that $\sigma_{\alpha\beta} = \sqrt{\sigma_{\alpha}^2\beta^2 + \sigma_{\beta}^2\alpha^2}$ where σ_{α}^2 is the squared standard error of the path coefficient between X and M , β^2 is the squared path coefficient between M and Y , σ_{β}^2 is the squared standard error of the path coefficient between M and Y , and α^2 is the squared path coefficient between X and M . Although the Sobel (1982) formula is still the most common method of calculating the standard error of the mediated effect (and the one that is automatically used to calculate the significance of the indirect effect in most SEM programs), MacKinnon et al. (2002) provide the formulas for and review the performance of several other standard error estimates. Researchers may wish to refer to this article for recommendations that are best suited to one’s precise study design. In addition to obtaining point estimates of the mediated effect, calculating the confidence intervals around the mediated effect is recommended (MacKinnon, Lockwood, & Williams, 2004). Further issues and recommendations related to the computation of the mediated effect and confidence intervals will be discussed in the next section on statistical power.

Statistical Power

Under many of the situations encountered in applied HIV/AIDS research, the mediated effects we seek can be small, and significant and/or clinically important mediated effects may be missed if the test used is underpowered. Although a complete discussion of power is beyond the scope of this paper; we will review this issue briefly and direct the reader to several excellent resources on this topic (e.g., Hoyle & Kenny, 1999; MacKinnon et al., 2002). For example, MacKinnon et al. (2002) review and compare the statistical power and Type I error rates of 14 approaches to mediation, all of which can be placed within the three broad classes already discussed (causal steps approaches, differences of coefficients methods, and product of coefficients methods). The authors observe that the conventional Baron and Kenny (1986) causal steps approach to mediation tends to have low power unless effect sizes and/or sample sizes are quite large. Although the product of coefficients methods tend to perform somewhat better, the commonly used Sobel (1982) standard error calculations also tend to suffer from low power.

One explanation for the low power of this test, however, may actually be due to multicollinearity between the independent variable and the mediator. As noted by Kenny and colleagues (Hoyle & Kenny, 1999; Kenny, Kashy, & Bolger, 1998), large α coefficients indicate that the independent variable has explained a great amount of variance in the mediator. There is subsequently little unique variance remaining in the mediator to predict the dependent variable, resulting in decreased precision of estimates of other coefficients and lowered power to test the mediated effect. Researchers should certainly be cognizant of the effect of a large correlation between the independent variable and the mediator on the power to test the mediated effect, and can refer to Hoyle and Kenny (1999) for more information regarding potential effects of collinearity on estimation of the coefficients and of the mediated effect.

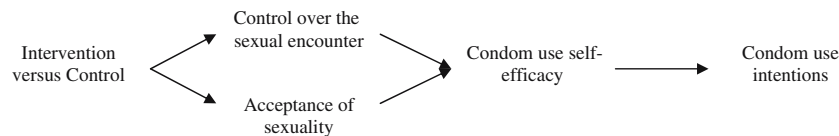
MacKinnon and colleagues have noted that the low power of the Sobel test may also stem from the fact that this test is based on a normal distribution, whereas the product of the coefficients rarely follows a normal distribution and is also associated with unbalanced confidence intervals (MacKinnon et al., 2002, 2004). Alternate methods of testing the significance of mediation, specifically resampling methods such as the bootstrap method (Efron, 1982), are being recommended because they have greater power to test the mediated effect and provide more accurate estimates of confidence intervals (MacKinnon et al., 2004;

Preacher & Hayes, 2004). Resampling methods are useful as a means to circumvent some of the standard assumptions about data (e.g., assumptions about the normality of variables, Diaconis & Efron, 1983). The logic behind these methods is that a large number of datasets (i.e., 1,000) can be generated by sampling, with replacement, from the original sample data. The generated datasets are then used to create an empirical sampling distribution that is used to test statistical hypotheses (i.e., values from the originally observed sample are compared to this empirical distribution, rather than to theoretical distributions, such as the t distribution).

Although resampling methods have long been recognized as useful, these methods were often overlooked because of intense computational and programming requirements. However, newer versions of computer programs allow for fast and simple applications of these methods. Resampling distributions can be created by using macros in standard statistical packages, such as SPSS and SAS (see Preacher & Hayes, 2004 for syntax examples). More importantly for this paper, resampling methods are implemented rather easily in a number of SEM packages (AMOS, EQS, Mplus). See Shrout and Bolger (2002) for a demonstration of the bootstrap method as applied to mediational analysis within an SEM framework.

Multiple-Mediator Models

It is rare that a single all-important mediator will completely account for the relationship between an independent and a dependent variable, particularly in models as complex as those typically used in HIV/AIDS research. Rather, researchers may wish to examine several mediating constructs within a single model, or to simultaneously examine individual components of a multi-component HIV prevention intervention program (c.f., West & Aiken, 1997). Analytically, models that allow for simultaneous tests to evaluate the relative contribution of each component will be the most useful and appropriate. As an example, Bryan et al. (1996) examined a single model that included the constructs of perceived benefits of condom use, control over the sexual encounter, acceptance of sexuality, attitudes toward condom use, and condom use self-efficacy as mediators of their intervention effects on intentions to use condoms. Further, their model included linkages among mediational variables, with self-efficacy predicted by control over the sexual encounter and acceptance of sexuality, and intention in turn predicted by self-efficacy (see diagram).



When multiple mediators or multiple independent or dependent measures (i.e., complex models) are tested, OLS regression is limited in that it is not possible to simultaneously examine multiple mediational chains, and an SEM approach is preferred. MacKinnon (2000) discusses issues related to calculating the significance of the mediated effect in multiple mediator models, and provides a demonstration of the contrasts necessary to compare the relative contribution of each mediator in multiple mediator models. Furthermore, SEM statistical packages are beginning to include capabilities to test for mediation in complex models by estimating not only the total indirect effect (Holbert & Stephenson, 2003), but the significance of each specific indirect effect. For example, in the diagram above one could estimate the indirect effect of the intervention on self-efficacy, specifically through control over the sexual encounter versus through acceptance of sexuality. Techniques for obtaining these significance levels in the SEM statistical packages Mplus and EQS and by hand are discussed in a later section.

Categorical Outcomes

Oftentimes in the field of HIV/AIDS research, we are interested in outcomes or mediators that do not distribute on a continuous scale, i.e., outcome of an HIV test (positive or negative), engagement in risky sexual behavior or not, sex with a risky partner or not, sharing injection equipment or not. As a specific example, it may sometimes be preferable to measure condom use at last intercourse (yes or no) in addition, or even as an alternative, to measuring an average of overall condom use (e.g., Bryan, Fisher, & Benziger, 2001). Here, we digress for a moment to clarify some terminology. Using the language of SEM, we distinguish between exogenous variables and endogenous variables. Exogenous variables are analogous to predictors in the regression context or independent variables in an experimental context, in that they are not regressed on any other construct, and the correlations among them are, by convention, always estimated. Endogenous variables are analogous to criterion or outcome variables in the regression context and mediators and dependent variables in the experimental context, and are regressed on one or more other variables in the model. Whether to estimate correlations among

endogenous variables is dictated by theory and empirical findings.

It is not appropriate to use OLS regression techniques when outcomes/endogenous variables are categorical, as two of the key assumptions underlying OLS regression analyses may be violated (the assumptions of normally distributed residuals and of homoscedasticity) and because the predicted values may be less than zero or greater than one, and thus theoretically inaccurate. Thus, when one has categorical endogenous variables (HIV test results, condom use at last intercourse, etc.), categorical regression techniques are recommended to avoid the biased or inaccurate results that may occur from analyzing categorical outcomes in an OLS framework (Cohen, Cohen, West, & Aiken, 2003). It is important to note that the necessity of using categorical regression techniques applies only to categorical *endogenous* variables; it is perfectly appropriate to estimate categorical *exogenous* variables (e.g., gender, marital status, race, intervention condition) within an OLS regression or SEM framework, provided the endogenous variables of interest are measured on a continuous scale.

The exact functions used to generate estimates in models with categorical endogenous variables are the probit and logistic functions. There are several distinctions between these two types of methods, although an extensive discussion of the differences between these functions is beyond the scope of this paper (see Agresti, 1990 or Hosmer & Lemeshow, 2000). For our purposes, it is important to know that either probit or logit estimation should provide reasonable estimates in simple mediational models with a single categorical outcome. For more complex models, the probit function has been the most commonly studied and utilized in the estimation of models with categorical endogenous variables in an SEM framework (cf., Muthén, 1979, 1983, 1984; Xie, 1989). Winship and Mare (1983) recommend using probit estimation for multiple-equation models with categorical endogenous variables and in particular in models that include both categorical and continuous endogenous variables. Mplus, the main SEM software program with capabilities to handle categorical data, includes the proper probit estimators for the techniques required for categorical outcomes using a weighted least squares estimator, and has recently

expanded to allow both probit and logistic estimation (see Mplus version 3). From a practical perspective and based on our review of the empirical and statistical literature (MacKinnon, *in press*; Winship & Mare, 1983), we recommend the use of probit estimation in complex models, but note that for simple models with a single categorical outcome, either logit or probit are appropriate. Also important is that calculation of the significance of the mediated effect becomes more complicated with categorical outcomes. MacKinnon and Dwyer (1993) provide an overview of the complexities of and potential solutions for testing the mediated effect in both probit and logistic models, and a forthcoming book by MacKinnon (*in press*) discusses this issue in even greater detail.

Multilevel or Clustered Data

Clustered data in which individuals are nested within groups (e.g., adolescents nested within schools, HIV+ patients nested within clinics) occur frequently in HIV/AIDS research, particularly in preventive intervention programs where the program is administered in groups. Clustered data violate the independence of observations assumption. Analyzing such data as individual observations (i.e., when patients from the same HIV clinic are treated as independent data points and clustering is ignored) can result in standard errors that are too small, *P*-values that are erroneously significant, and conclusions which may be wrong (Bland, 2004; Puffer et al., 2003). Thus, analytic procedures must be used that account for non-independence of observations, and some SEM packages (EQS and Mplus at the time of this writing) now include estimation procedures that can handle such “multilevel” data. Krull and MacKinnon (1999, 2001) examined mediation models within a multilevel framework, showing that both the product of coefficients and the difference of coefficients methods provided reasonable estimates of the mediated effect when translated to multilevel models. The authors indicated a slight preference for the product of coefficients method, as this method performed somewhat better and yields a greater degree of information than the difference of coefficients method.

Data measured longitudinally are a special case of multilevel data in that time points are clustered within individuals. Longitudinal designs may be preferred in HIV/AIDS research due to the types of theoretical questions that can be answered (e.g., changes in risk behavior over time) by measuring individuals repeatedly over a specified period of time, including questions about the way mediators function over time.

Two excellent resources on mediation with longitudinal data are available for researchers: Cheong, MacKinnon, and Khoo (2003) describe and give an example of a parallel-process method for conducting longitudinal mediational analyses within an SEM framework, in which the growth in the mediator over time is linked to the growth in the outcome over time; and Cole and Maxwell (2003) provide an overview of common issues that arise when testing for mediation with longitudinal data, including strategies for overcoming common obstacles and for developing appropriate models.

Nonnormal and Missing Data

A further advantage of examining mediating variables within an SEM framework is the ease with which problems in the data, such as nonnormality or missing data, can be handled using special and/or robust estimators that can be implemented in most SEM packages. Nonnormal data (i.e., excessive skew and/or kurtosis) can be addressed using the Satorra-Bentler robust maximum likelihood estimator (Satorra & Bentler, 1988, 1994). Depending on the software package that is used, data that are either missing completely at random (MCAR) or missing at random (MAR) can be addressed with the full information (direct) maximum likelihood estimator (Arbuckle, 1996) or with an Expectation Maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977). Recently, estimators have been developed that are capable of addressing data that are both nonnormal and missing (Yuan & Bentler, 2000).

Interpretation of SEM Models

We turn now to issues of the interpretation of path analytic mediational models within an SEM context. This includes interpreting what the data from the estimation of such a model tell us conceptually and theoretically, as well as understanding when a model is “good” in the sense that it is statistically valid.

Fit of Individual Paths

Situations may occur where the overall model demonstrates good fit, but model relationships are weak or not meaningful (MacCallum & Austin, 2000). It is thus critical to first examine the magnitude and significance of each model path, as well as the proportion of variance accounted for in each endogenous variable by the set of exogenous variables. All SEM software packages

provide several key pieces of information to evaluate the performance of specific paths: an unstandardized path coefficient, a standard error for each path coefficient, a test of significance of each coefficient (evaluated along a z -distribution), and some form of a standardized estimate for each coefficient (note that the precise standardized estimates that are reported differ by the statistical package used). It is these path coefficients that are in large measure what we use to determine what the model is “telling us” conceptually and theoretically. Did the intervention impact perceived susceptibility to HIV? Does finding out one’s HIV serostatus impact risk behavior directly or indirectly? Are men or women more likely to have an internal locus of control and does locus of control influence adherence to antiretroviral regimen? All of these questions are answered by the examination of the path coefficients in a good-fitting SEM analysis.

Overall Model Fit

The next step in determining the statistical validity of a model is to assess fit. Overall fit of SEM models can be determined by a multitude of fit indices, each with associated advantages and disadvantages. Hu and Bentler (1998, 1999) provide an overview of available indices, as well as guidelines for assessing the acceptable range for each index. The authors recommend reporting both the chi-square (χ^2) value and the Standardized Root Mean Square Residual (SRMR; Bentler, 1995), supplemented by one of several other indices, such as the Root Mean Square Error of Approximation (RMSEA; Steiger & Lind, 1980) or the Comparative Fit Index (CFI; Bentler, 1990). Though we recommend reporting the χ^2 value for completeness, one should exercise caution in basing decisions about the adequacy of fit on χ^2 values. The χ^2 is a test of the null hypothesis which, in the case of SEM, is that the model *fits* the data. Thus we are in the unusual case of wanting to *accept* the null. As with any test of the null hypothesis, the power of the χ^2 test is hugely dependent upon sample size. A model estimated with a relatively small sample will almost always yield a non-significant χ^2 , while a model estimated with a large sample size will almost always yield a significant χ^2 , regardless of the actual fit of the model. It is for this reason that it is crucial to report and to base decisions about fit on indices that are less affected by sample size. We recommend the SRMR, the CFI, and the RMSEA, as each has desirable theoretical and statistical properties under most circumstances, and can be used in conjunction with one another to assess overall fit. General

guidelines for specific cut-off points in assessing overall model fit are values close to .08 or lower for SRMR, close to .95 or higher for CFI, and close to .06 or lower for RMSEA are indicative of adequate fit. We caution the reader, however, that there are ongoing debates in the methodological literature concerning which fit indices are most appropriate, what the cut-offs should be, and whether cut-offs should even be recommended at all (e.g., Fan & Sivo, 2005; Marsh, Hau, & Wen, 2004).

A further caveat in interpreting overall model fit is that it is possible to have a good-fitting, but theoretically and conceptually uninformative model if researchers are not parsimonious in specifying model relationships. If there are as many parameters estimated as there are degrees of freedom, the model will be saturated, leading to a “perfect” fit of the model to the data. One certain way of determining that a model may have too many parameters included is if there are no degrees of freedom left to test fit. Certainly in some situations this is fine (e.g., a confirmatory factor analysis with three indicators), but in most of the complex models we test in HIV/AIDS research, zero degrees of freedom is indicative of a saturated model and a lack of parsimony in specifying model relationships. Short of having zero or few degrees of freedom, diagnosing overfit is somewhat subjective. When estimating a model that requires continually adding path after path in order to get fit indices to the magical cut-points, it is likely that the model suffers from overfit, which compromises its interpretability, parsimony, and replicability.

Programs for Conducting SEM Analyses

There are a number of computer software packages available for estimating structural equation models. Though by no means a comprehensive list, some of the most commonly used packages in the social sciences are AMOS (Arbuckle, 2003), PROC CALIS, which is included as part of the broader SAS software (SAS Institute, 2005), EQS (Bentler, 1995), LISREL (Jöreskog & Sörbom, 2004), MPlus (Muthén & Muthén, 1998–2005), and Mx (Neale, Boker, Xie, & Maes, 2003). In the majority of situations (e.g., simple models with normal and non-missing data), any one of these packages will adequately estimate appropriate models. In special situations, some programs have advantages over others, and the specific programs that will be the most useful will depend on the nature of the data. Researchers should thus verify a program’s capabilities by examining the user’s manual before conducting the desired analyses. Given

that EQS and Mplus are quite flexible in addressing a variety of complex data situations, we will demonstrate the use of these two programs further by applying them in our examples. Both of these programs are equation-based (though EQS does have a graphical interface), in that one essentially specifies the multiple regression equations being simultaneously estimated in the overall SEM model. It is our bias to prefer equation-based (over graphical interface-based) programs as they force the user to completely specify the equations, error structure, and correlations to be estimated.

Examples Demonstrating Tests of Mediation in an SEM Framework

Theoretical Model Testing

Within the context of a larger study testing the distal effect of alcohol on condom use, Theory of Planned Behavior (TPB) constructs were measured longitudinally in a sample of 300 adolescents involved with the Denver metro area juvenile justice system. For in-depth analysis, see Bryan, Rocheleau, Robbins, and Hutchison (2005). Here, we estimated the basic TPB model in EQS (see syntax in Appendix A), with attitudes, norms, and self-efficacy predicting intentions, which in turn predicted behavior six months later. Consistent with the TPB, we also tested the direct effect of self-efficacy on behavior. This model was estimated using the 227 adolescents who had complete data, using listwise deletion for missing data². Univariate skewness values ranged from $-.07$ to -1.12 and univariate kurtosis values ranged from $-.79$ to $.91$, confirming that the variables were indeed normally distributed (West, Finch, & Curran, 1995) and that no special estimators to address nonnormality were necessary.

The overall fit indices and parameter estimates, taken directly from EQS output, are shown in Table 1. EQS presents results in the form of regression equations, where the first set of equations is the unstandardized parameter estimates, and the second set is the standardized parameter estimates. The number directly under the unstandardized parameter estimate is the standard error of the parameter, and the number

directly under the standard error is the test-statistic (distributed as z). As shown, the hypothesized relationships were supported, such that attitudes, norms, and self-efficacy all predicted intentions to use condoms ($\beta = .16$, $P < .01$ for attitudes; $\beta = .25$, $P < .01$ for norms, and $\beta = .30$, $P < .01$ for self-efficacy), while both intentions and self efficacy predicted behavior ($\beta = .42$, $P < .01$ and $\beta = .14$, $P < .05$, respectively). The model accounted for 29% of the variance in intentions, and 25% of the variance in behavior. However, the overall fit of the model to the data was only moderate, $\chi^2(2, N = 227) = 21.70$, $P < .01$; CFI = .91; RMSEA = .21, with 90% confidence intervals (CI) = .14–.29; SRMR = .07, indicating some room for improvement. Note in particular the high value of the RMSEA.

An exploratory approach was taken to determine whether model fit could be improved. EQS provides LaGrange Multiplier (LM) values (termed “modification indices” in some programs, such as LISREL and Mplus) that approximate how the value of χ^2 might improve (i.e., decrease) for each possible path or relationship that could be incorporated into the model (Chou & Bentler, 1990). Table 2 depicts the EQS output of the LM tests for adding parameters. As shown by this output, including a path between V2 and V6 (attitudes and behavior) would result in an estimated decrease in the χ^2 value of 17.23 and a standardized parameter estimate of .47. A path from attitudes to behavior was included in the model, resulting in improved overall model fit, $\chi^2(1, N = 227) = 3.77$, $P = .05$; CFI = .99; RMSEA = .11, 90% CI = .00–.24; SRMR = .025. A one degree of freedom χ^2 change test (χ^2_{Δ}) can also be computed that can directly compare nested models and test the significance of the change in χ^2 occasioned by the freeing of an additional parameter. In this case, $\chi^2_{\Delta}(1, N = 227) = 17.93$, $P < .01$, indicating a significant improvement in fit. The syntax that generated this final model appears in Appendix A. The standard convention is to report the standardized parameter estimates and indications of significance ($*P < .05$, $**P < .01$) in the figures presenting the models, and this is what has been done in Fig. 2. The coefficients in Fig. 2 are from the final model including the path from attitudes to behavior (dotted in the figure).

Although examination of the LM tests (or modification indices) can be useful in identifying sources of model misfit, we recommend that researchers exercise a great deal of caution in using these tests (see MacCallum, 1986 for a full discussion of the limitations). A key limitation of this approach is that LaGrange Multiplier tests are exploratory and are guided by data,

² It is also possible to use the EM algorithm to obtain maximum likelihood estimates for missing data in EQS. This procedure makes the EQS programming slightly more complex, so in this example we chose to use listwise deletion as it keeps the programming language more simplified. For sample programs using the EM algorithm for missing data in EQS, we refer the reader to the EQS program manual (Bentler, 1995).

Table 1 Output from EQS demonstrating overall fit and significance of model parameters of theoretical model

CHI-SQUARE = 21.695 BASED ON 2 DEGREES OF FREEDOM
 PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS .00002
 THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS 20.686.

FIT INDICES

BENTLER-BONETT NORMED FIT INDEX = .902
 BENTLER-BONETT NON-NORMED FIT INDEX = .534
 COMPARATIVE FIT INDEX (CFI) = .907
 BOLLEN (IFI) FIT INDEX = .910
 MCDONALD (MFI) FIT INDEX = .958
 LISREL GFI FIT INDEX = .965
 LISREL AGFI FIT INDEX = .735
 ROOT MEAN-SQUARE RESIDUAL (RMR) = .055
 STANDARDIZED RMR = .067
 ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA) = .209
 90% CONFIDENCE INTERVAL OF RMSEA (.135, .291)

MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @

INTENT =V5 = .163*V2 + .241*V3 + .421*V4 + 1.000 E5
 .061 .058 .085
 2.683@ 4.178@ 4.984@

BEHAVIOR= V6 = .620*V5 + .282*V4 + 1.000 E6
 .095 .132
 6.556@ 2.142@

STANDARDIZED SOLUTION: R-SQUARED

INTENT =V5 = .164*V2 + .253*V3 + .303*V4 + .844 E5 .287
 BEHAVIOR= V6 = .420*V5 + .137*V4 + .869 E6 .245

Table 2 EQS output demonstrating LaGrange multipliers and indirect effects of theoretical model

LAGRANGE MULTIPLIER TEST (FOR ADDING PARAMETERS)
 ORDERED UNIVARIATE TEST STATISTICS:

NO	CODE	PARAMETER	CHI-SQUARE	PROB.	HANCOCK 2 DF PROB.	PARAMETER CHANGE	STANDARDIZED CHANGE
1	2 11	V6,V2	17.233	.000	.000	.383	.471
2	2 11	V6,V3	6.670	.010	.036	.231	.273

DECOMPOSITION OF EFFECTS WITH NONSTANDARDIZED VALUES
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
 PARAMETER INDIRECT EFFECTS

BEHAVIOR = V6 = .086*V2 + .126 V3 + .221*V4 + .525 E5
 .035 .038 .059 .094
 2.420@ 3.350@ 3.724@ 5.604@

DECOMPOSITION OF EFFECTS WITH STANDARDIZED VALUES
 PARAMETER INDIRECT EFFECTS

BEHAVIOR = V6 = .058*V2 + .090 V3 + .108*V4 + .300 E5

rather than theory, and the model changes offered by the LM tests might not be conceptually appropriate or even realistic (e.g., LM tests could conceivably recommend paths that go backward in time). Furthermore, the inclusion of additional paths may change previously observed model relationships. In the above example, the direct path from self-efficacy to behavior was no longer significant with the inclusion of the direct path from attitudes to behavior. When using LM tests, researchers should critically evaluate whether the suggested path is appropriate, should offer a theoretical justification for the inclusion of the paths, and

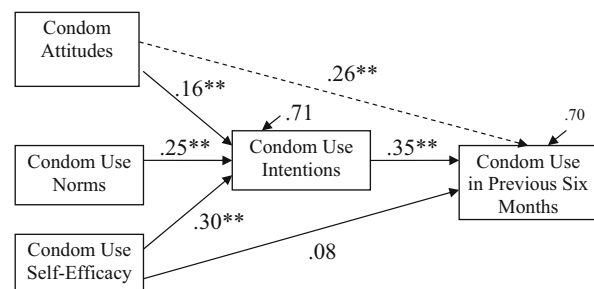


Fig. 2 Demonstration of mediational analyses in an SEM framework to test a theoretical model. Coefficients are standardized path coefficients. * $P < .05$, ** $P < .01$

should replicate all results in an independent sample of participants.

Once a final model is settled upon, one can ask any number of mediational questions. As an example, let us focus on the relationship of attitudes towards condom use to subsequent condom use behavior. One question is whether there was a direct effect of attitudes on behavior, a question easily answered by the significant parameter estimate for the path from attitudes to behavior in the overall model. A second question is whether there is any additional effect of attitudes on behavior that is mediated through intentions. In other words, what is the indirect effect of attitudes on behavior? In this case printing the decomposition of effects (i.e., total, direct, and indirect) in the output simply requires “effects = yes;” under the “PRINT” subcommand in EQS (see Appendix A). This output, as can be seen in Table 2, displays all the tests of the indirect effects and includes the EQS implementation of the Sobel (1982) test for significance. The indirect effect of attitudes (V2) on behavior (V6) carries an unstandardized parameter estimate of .086, a standard error of .035, and a Sobel test of 2.42, which is significant at $P < .05$. The significance of each indirect effect is calculated using the product of coefficients as the numerator ($\alpha\beta$) and Sobel’s (1982) standard error as the denominator ($\sqrt{\sigma_{\alpha}^2\beta^2 + \sigma_{\beta}^2\alpha^2}$). This same number can be arrived at by hand. The indirect effect of attitudes on behavior, through intentions, is computed by dividing the unstandardized mediated effect parameter ($\alpha\beta$ or $.163 \times .525 = .0856$) by Sobel’s (1982) standard error estimate ($\sqrt{.061^2 \times .525^2 + .093^2 \times .163^2}$). Doing so yields the same estimate of the mediated effect and its significance ($.0856 / .035$; $z = 2.42$, $P < .05$) provided in the EQS output. The 95% lower and upper confidence limits around the mediated effect were .017 and .154, respectively (obtained by the following formula: $\alpha\beta \pm z_{\text{Type I Error}} [\sigma_{\alpha\beta}]$ or $.0856 \pm 1.96 [.035]$). We could thus conclude from these findings that there are both significant direct and indirect (through intentions) effects of attitudes on condom use behavior.

Examination of Multi-Component Interventions

Data from an ongoing randomized controlled trial of a sexual risk reduction intervention administered to incarcerated adolescents will be used to demonstrate the use of mediational analysis in the evaluation of individual components of a multi-component intervention program. Although data collection on this project is still active with a planned sample size of 480, the results presented here are based on the total

number of participants to date who have completed the pretest and immediate posttest measures and who were eligible to complete the three month behavioral follow-up ($n = 251$).

Participants were randomly assigned to receive one of three intervention programs in groups of no more than ten: an information-only control condition ($n = 87$), a sexual risk reduction condition ($n = 101$), or a combined sexual risk and alcohol risk reduction condition ($n = 63$). For ease of illustration, however, we combine the two intervention programs into a single condition, resulting in two levels of the program condition variable (0 = control condition; 1 = either of the two intervention conditions). The analyses presented here are based on immediate posttest (occurring directly after administration of the program) assessments of attitudes towards condom use, norms for condom use, self-efficacy for condom use, and intentions to use condoms in the future, as well as condom use behavior measured three months after program administration. Due to the high number of participants who were missing data at the 3 month follow-up ($n = 130$ of 251), the model was estimated in Mplus using the full information maximum likelihood (FIML) estimator. The variables were normally distributed with univariate skewness values ranging from -1.0 to -1.31 and univariate kurtosis values ranging from -0.89 to 1.43 .

The estimated model is depicted in Fig. 3, and example Mplus output of model parameters is depicted in Table 3. The syntax used to generate this model appears in Appendix B. As shown in Table 3, Mplus provides five columns of output for each estimated parameter. The first column depicts the unstandardized coefficient; the second column, the standard error; the third column, the z -score demonstrating the significance of each parameter; the fourth column, the standardized solution; and, the fifth column, the completely standardized solution in which both endogenous and exogenous variables are standardized. The completely standardized coefficients (reported in Fig. 3) should be presented across the majority of situations.

As can be seen from both Fig. 3 and Table 3, there was a significant effect of condition on both attitudes and self-efficacy ($\beta = .14$, $P < .05$ and $\beta = .22$, $P < .01$, respectively), but not on norms ($\beta = .07$, *n.s.*). The direct effect of condition on intentions was not significant with the three psychosocial constructs included in the model ($\beta = .06$, *n.s.*). Self-efficacy and norms, but not attitudes, significantly predicted intentions to use condoms (self-efficacy, $\beta = .37$, $P < .01$; norms, $\beta = .24$, $P < .01$; and, attitudes, $\beta = .03$, *n.s.*). Intentions, in turn, significantly predicted condom use behavior at the 3 month follow-up ($\beta = .72$, $P < .01$).

Fig. 3 Demonstration of mediational analyses in an SEM framework to test individual components of multi-component interventions. Coefficients are standardized path coefficients. * $P < .05$, ** $P < .01$

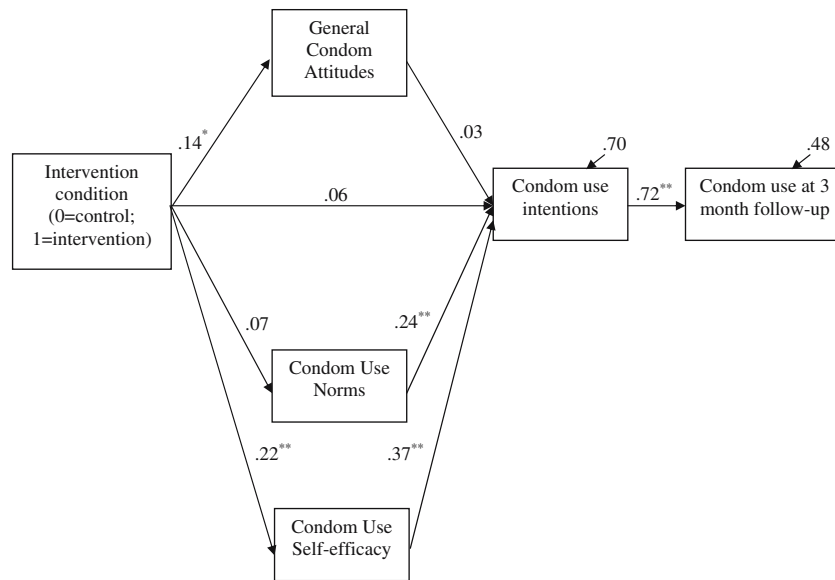


Table 3 Output from Mplus depicting parameter estimates of multicomponent intervention

	Estimates	S.E.	Est./S.E.	Std	StdYX
MODEL RESULTS					
ATTITUDE ON CONDITION	12.056	5.602	2.152	12.056	0.142
SELFEFF ON CONDITION	17.803	5.207	3.419	17.803	0.223
NORMS ON CONDITION	12.877	11.667	1.104	12.877	0.074
INTENT ON					
ATTITUDE	0.065	0.138	0.471	0.065	0.032
SELFEFF	0.783	0.151	5.175	0.783	0.365
NORMS	0.238	0.061	3.911	0.238	0.242
CONDITION	10.227	9.820	1.041	10.227	0.060
BEHAVIOR ON INTENT	1.442	0.207	6.974	1.442	0.720
ATTITUDE WITH					
SELFEFF	0.083	0.011	7.494	0.083	0.542
NORMS	0.125	0.023	5.395	0.125	0.374
SELFEFF WITH NORMS	0.130	0.022	5.965	0.130	0.412

Table 4 Mplus output for specific indirect effects and confidence limits around the indirect effects

	Lower .5%	Lower 2.5%	Estimates	Upper 2.5%	Upper .5%
Effects from CONDTIO to INTENT					
Sum of indirect			17.786	17.786	0.104
Specific indirect					
INTENT ATTITUDE CONDTIO			0.783	0.783	0.005
INTENT SELFEFF CONDTIO			13.936	13.936	0.081
INTENT NORMS CONDTIO			3.067	3.067	0.018
Effects from CONDTIO to BEHAVIOR					
Specific indirect					
BEHAVIOR INTENT SELFEFF CONDTIO			20.100	20.100	0.058
	Lower .5%	Lower 2.5%	Estimates	Upper 2.5%	Upper .5%
Effects from CONDTIO to INTENT					
Specific indirect					
INTENT ATTITUDE CONDTIO	-3.603	-2.555	0.783	4.121	5.169
INTENT SELFEFF CONDTIO	1.328	4.342	13.936	23.529	26.543
INTENT NORMS CONDTIO	-4.385	-2.604	3.067	8.739	10.520
Effects from CONDTIO to BEHAVIOR					
BEHAVIOR INTENT SELFEFF CONDTIO	0.509	5.193	20.100	35.007	39.691

There was good fit of the model to the data, $\chi^2(4, N = 251) = 1.69, n.s.$; CFI = 1.00; RMSEA = .00, with 90% CI from .00 to .06; SRMR = .02. The model accounted for 30% of the variance in intentions, and 52% of the variance in behavior. Although not depicted in Fig. 3 for ease of demonstration, Table 3 shows that the disturbance terms (errors in prediction) were correlated among each psychosocial construct. Direct paths from program condition and the three psychosocial constructs to behavior were estimated to examine whether there were remaining direct effects of these constructs on behavior. None of these paths were significant and were thus not included in the final model.

Mplus calculates each specific indirect effect using a simple line of syntax (see the “model indirect” statement in Appendix B), allowing easy examination of the significance of the mediated effect. The Mplus output of each specific indirect effect and the 90% and 95% confidence intervals around those effects are shown in Table 4. For example, the specific indirect effect of condition on intentions, through self-efficacy, is given by the parameter estimate of 13.936, and the standard error estimate of 4.895. The Sobel test is then $13.936/4.895 = 2.847, P < .01$ with 95% lower and upper confidence limits of 4.34 and 23.53, as can be seen in the Mplus output. One could conclude from this analysis that there was a significant indirect (mediated) effect of the intervention on intention through self-efficacy. It is also important to note that this analysis demonstrates that neither attitudes nor norms significantly mediated the relationship between condition and intentions, which is not surprising, given that the intervention had no effect on norms for condom use as compared to the information-only control condition, and that attitudes were unrelated to intentions.

A feature that is, at the time of this writing, unique to Mplus is that it is possible to calculate the significance of the specific indirect effect when there are multiple mediational pathways in a model. We already demonstrated how to calculate the significance of the indirect path from condition to *intentions* via self-efficacy using Mplus and by hand. However, intervention condition may also relate indirectly to *behavior* through the psychosocial constructs and intentions. We demonstrate the syntax for calculating the significance of the specific indirect effect from condition to behavior via self-efficacy and intentions in Appendix B (“behavior ind intent selfeff condition”). As shown in Table 4, this specific indirect effect was statistically significant ($z = 2.64, P < .01$), with unstandardized parameter estimate of 20.10 and a standard error of 7.61. For users of other programs, the parameter

estimates and standard errors of specific indirect effects must be calculated by hand (see Fox, 1980, 1985).

Special Considerations in the use of Mediation via Path Analysis

Oftentimes researchers will not be interested in testing a single model, but will instead propose, a priori, competing models in order to determine how well the models compare to one another. Model comparison strategies differ depending on whether the models to be compared are nested or non-nested. Nested models refer to situations in which one model can “fit inside” another model. In other words, if constraints can be made on a model that result in a simpler model, the simpler model is nested within the more complex model. Take the theoretical TPB model we estimated above as an example—here, the first model estimated (the model without a direct path from attitudes to behavior) was nested within the final model because the initial, simpler model treated the path from attitudes to behavior as constrained to zero, while the more complex model freely estimated the path. Nested models can be compared directly through the use of χ^2 difference tests, as we did above in the theoretical example. If the χ^2 value is significant, the more complex model should be retained; non-significant χ^2 values imply that there is no statistical difference between models, meaning that the simpler (more parsimonious) model should be retained. Non-nested models cannot be compared in such a direct manner. Instead, researchers must rely on overall goodness-of-fit statistics, the magnitude and significance of the path coefficients, and the theoretical justification for each model for model comparison. Furthermore, several information criterion measures can be used (Haughton, Oud, & Jansen, 1997), such as Akaike’s information criterion (AIC) and the Bayesian information criterion (BIC), in which smaller values indicate better-fitting models.

SEM output should always be examined critically for several key pieces of information to ensure that the model estimation “converged”, i.e., reached an appropriate solution. The output can be used to verify that the model was estimated as intended, and that all parameter estimates and standard errors are within an appropriate range. Technical output can be examined to determine whether the model successfully converged and the number of iterations that were required to do so. Most programs give warning messages if the model estimation did not terminate normally, although sometimes information on model fit and the signifi-

cance of path coefficients is provided even when the model was in fact misspecified or inaccurate (as indicated by out-of-range parameter estimates such as negative variances).

SEM models may either be recursive or non-recursive. In recursive models, the effects of one variable on another are directional, flowing only one way (correlations among exogenous constructs are still allowed). Non-recursive models may include bidirectional paths, loops among variables, and correlated disturbances among endogenous variables. While non-recursive models are in some cases plausible (e.g., correlations among the disturbances of mediators in multiple-mediator models), they tend to have estimation problems and in many cases are simply inestimable. For example, if one believes that intentions are related to current behavior and then behavior, in a feedback loop, also affects intentions, this is certainly logically possible. But in an SEM context, such a model would not be estimable because that particular section of the model would be under-identified. We advise extreme caution in the estimation of models including non-recursivity in most circumstances.

Conclusions

It is our hope that this paper has “demystified” the use of the SEM framework for mediational analysis for those unfamiliar with the techniques involved, provided updated information for those who are current users, and inspired those who would like to expand their use of these techniques to data sets and questions that are more complex than what they might be used to. As we noted at the outset of this paper, the class of analyses available under SEM to test mediational questions in HIV/AIDS research and other domains is incredibly flexible and powerful. From hypothesis generation to theory testing to probing intervention effects, the possibilities for mediational analysis via SEM are expanding due to the addition of accessible statistical packages that are continually incorporating additional capabilities and estimators to adapt to complex research designs. As long as researchers are aware of the underlying assumptions and appropriate use of these techniques, we see them as having an enormous capacity to advance theorizing on behavior change generally, and theoretical aspects of HIV/AIDS research specifically. Importantly, they allow us to understand which aspects of our interventions are successful and which are not. The combination of these two research questions, which are entirely amenable to

tests of mediation via SEM techniques, has the capacity to allow us to design and evaluate HIV/AIDS risk reduction interventions that are optimally effective.

Acknowledgement The authors would like to thank Dr. David Kenny for serving as a reviewer on this paper. His contributions were invaluable. Preparation of this paper was supported by grants from the National Institute on Alcoholism and Alcohol Abuse (AA013844-01) and the National Institute on Drug Abuse (DA019139-01) awarded to the first author.

Appendix

Appendix A

EQS Syntax

/TITLE

Theory-testing model in EQS for Bryan, Schmiede,
& Broaddus

/SPECIFICATIONS

DATA = 'c:\eqs files\data\psex.ess';

VARIABLES = 6; CASES = 300;

METHOD = ML; ANALYSIS = COVARIANCE;

MATRIX = RAW;

/LABELS

v1 = caseid; v2 = attitude; v3 = norms; v4 = selfeff;
v5 = intent; v6 = behavior;

/EQUATIONS

v6 = *v4 + *v5 + *v2 + e6;

v5 = *v2 + *v3 + *v4 + e5;

/VARIANCES

e6 = *; e5 = *;

/COVARIANCES

v2,v3 = *; v2,v4 = *; v3,v4 = *;

/TEC

iter = 500;

/PRINT

fit = all;

correlations = yes;

effects = yes;

/WTEST;

/LMTEST;

/END

Appendix B

Mplus Syntax

Title: Intervention model for Bryan, Schmiede, & Broaddus;

Data:
 File is c:\intex;
 Format is 1f7.0 8f7.2;
 Variable:
 Names are
 case condition attitude selfeff norms intent intalc
 t3intent behavior;
 Usevariables are
 condition attitude selfeff norms intent behavior;
 Missing are all .;
 Analysis:
 Type = general missing h1;
 Estimator is ML;
 Model:
 attitude on condition;
 selfeff on condition;
 norms on condition;
 intent on attitude;
 intent on selfeff;
 intent on norms;
 intent on condition;
 attitude with selfeff;
 attitude with norms;
 selfeff with norms;
 behavior on intent;
 Model indirect:
 intent ind attitude condition;
 intent ind selfeff condition;
 intent ind norms condition;
 behavior ind intent selfeff condition;
 Output:
 Standardized Sampstat Residual Modindices(4)
 Patterns H1se Cinterval;

References

- Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.
- Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques*. Mahwah, NJ: Erlbaum.
- Arbuckle, J. L. (2003). *Amos 5.0 update to the amos user's guide*. Chicago IL: Smallwaters Corporation.
- Albarracin, D., Fishbein, M., & Middlestadt, S. (1998). Generalizing behavioral findings across times, samples, and measures: A study of condom use. *Journal of Applied Social Psychology, 28*(8), 657–674.
- Albarracin, D., & Wyer, R. S. (2000). The cognitive impact of past behavior: Influences in beliefs, attitudes, and future behavioral decisions. *Journal of Personality and Social Psychology, 79*, 5–22.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality & Social Psychology, 51*, 1173–1182.
- Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology, 31*, 419–456.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.
- Bland, J. M. (2004). Cluster randomised trials in the medical literature: Two bibliometric surveys. *BMC Medical Research Methodology, 4*(21) [electronic resource].
- Bollen, K. A. (1987). Total, direct, and indirect effects in structural equation models. In C. C. Clogg (Ed.), *Sociological methodology 1987* Washington DC: American Sociological Association.
- Bollen, K. A. (1989). *Structural equations with latent variables*. Oxford, England: John Wiley & Sons.
- Bryan, A. D., Aiken, L. S., & West, S. G. (1996). Increasing condom use: Evaluation of a theory-based intervention to decrease sexually transmitted disease in women. *Health Psychology, 15*, 371–382.
- Bryan, A. D., Fisher, J. D., Fisher, W. A., & Murray, D. M. (2000). Understanding condom use among heroin addicts in methadone maintenance using the information-motivation-behavioral skills model. *Substance Use and Misuse, 35*, 451–471.
- Bryan, A. D., Fisher, J. D., & Benziger, T. J. (2001). Determinants of HIV risk behavior among Indian truck drivers. *Social Science and Medicine, 53*, 1413–1426.
- Bryan, A., Rocheleau, C. A., Robbins, R. N., & Hutchison, K. E. (2005). Condom use among high-risk adolescents: Testing the influence of alcohol use on the relationship of cognitive correlates of behavior. *Health Psychology, 24*, 133–142.
- Cheong, J., MacKinnon, D. P., & Khoo, S. -T. (2003). Investigation of mediational processes using parallel process latent growth curve modeling. *Structural Equation Modeling, 10*(2), 238–262.
- Chou, C. -P., & Bentler, P. M. (1990). Model modification in covariance structure modeling: A comparison among likelihood ratio, Lagrange Multiplier, and Wald tests. *Multivariate Behavior Research, 25*(1), 115–136.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*, 558–577.
- Collins, L. M., Graham, J. W., & Flaherty, B. P. (1998). An alternative framework for defining mediation. *Multivariate Behavioral Research, 33*, 295–312.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum-likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39*, 1–38.
- Diaconis, P., & Efron B. (1983). Computer intensive methods in statistics. *Scientific American, 248*, 116–130.
- Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. *Society for Industrial and Applied Mathematics CBMS-NSF Regional Conference Series (No. 38)*.
- Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling, 12*(3), 343–367.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

- Fisher, J. D., & Fisher, W. A. (2002). The information-motivation-behavioral skills model. In R. DiClemente R. Crosby & R. Kegler (Eds.), *Emerging promotion research and practice* (pp. 40–70). San Francisco, CA: Jossey Bass.
- Fox, J. (1980). Effect analysis in structural equation models: Extensions and simplified methods of computation. *Sociological Methods and Research*, 9, 3–28.
- Fox, J. (1985). Effect analysis in structural-equation models II: Calculation of specific indirect effects. *Sociological Methods and Research*, 14, 81–95.
- Houghton, D. M. A., Oud, J. H. L., & Jansen, R. A. R. G. (1997). Information and other criteria in structural equation model selection. *Communications in Statistics: Simulation and Computation*, 26, 1477–1516.
- Holbert, R. L., & Stephenson, M. T. (2003). The importance of indirect effects in media effects research: Testing for mediation in structural equation modeling. *Journal of Broadcasting & Electronic Media*, 47(4), 556–572.
- Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, 65, 599–610.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York: Wiley.
- Hoyle, R. H. (1995). *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, CA: Sage.
- Hoyle, R. H., & Kenny, D. A. (1999). Sample size, reliability, and tests of statistical mediation. In: Hoyle R. H. (Ed). *Statistical strategies for small sample research* (pp. 195–222). Thousand Oaks, CA: Sage Publications.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424–453.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Jöreskog, K. G., & Sörbom, D. (2004). *LISREL 8.7 for Windows [Computer Software]*. Lincolnwood, IL: Scientific Software International, Inc.
- Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*, 117, 348–357.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, 5, 602–619.
- Kenny, D. A. (2006). Mediation. <http://davidakenny.net/cm/mediate.htm>. Obtained January 26, 2006.
- Kenny D. A., Kashy D. A., & Bolger N. (1998). Data analysis in social psychology. In: Gilbert D. T. Fiske S. T. & Lindzey G. (Eds). *The handbook of social psychology vol. 1*, (4th ed.) (pp. 233–265). New York, NY, US: McGraw-Hill.
- Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, 59, 877–883.
- Krull, J. L., & MacKinnon, D. P. (1999) Multilevel mediation modeling in group-based intervention studies. *Evaluation Review*, 23, 418–444.
- Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, 36, 249–277.
- Loehlin, J. C. (1992). *Latent variable models: An introduction to factor, path, and structural analysis*. (2nd Ed.). Hillsdale, NJ: Erlbaum.
- MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, 100(1), 107–120.
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, 51, 201–226.
- MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, 114, 185–199.
- MacKinnon, D.P. (in press). *Introduction to statistical mediation analysis*, Mahwah, NJ: Erlbaum.
- MacKinnon, D. P. (1994). Analysis of mediating variables in prevention, intervention research. In: A. Cazares, & L.A. Beatty (Eds.), *Scientific methods for prevention intervention research* (pp. 127–153). Washington, DC: NIDA Research Monograph 139, DHHS Pub. 94–3631.
- MacKinnon, D. P. (2000). Contrasts in multiple mediator models. In: J. Rose, L. Chassin, C. C. Presson, & S. J. Sherman (Eds.), *Multivariate applications in substance use research: New methods for new questions* (pp. 141–160). Mahwah, NJ: Erlbaum.
- MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, 17, 144–158.
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7, 83–104.
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39, 99–128.
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science*, 1, 173–181.
- MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, 30, 41–62.
- Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11(3), 320–341.
- Marcoulides, G. A., & Schumacker, R. E. (1996). Introduction. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques*. (pp. 1–6). Mahwah, NJ: Erlbaum.
- Muthén, B. O. (1979). A structural probit model with latent variables. *Journal of the American Statistical Association*, 74, 807–811.
- Muthén, B. O. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics*, 22, 43–65.
- Muthén, B. O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49, 115–132.
- Muthén, L. K., & Muthén, B. O. (1998–2005). *Mplus User's Guide*. (3rd ed.). Los Angeles, CA: Muthén & Muthén.
- Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical Modeling*. VCU Box 900126, Richmond, VA 23298: Department of Psychiatry. 6th Edition.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, U.K.: Cambridge University Press.
- Preacher, K. J. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36(4), 717–731.

- Puffer, S., Torgerson, D. J., & Watson, J. (2003). Evidence for risk of bias in cluster randomized trials: Review of recent trials published in three general medical journals. *BMJ: British Medical Journal*, *327*(7418), 785–789.
- Rosenstock, I. M. (1990). The health belief model: Explaining health behavior through expectancies. In K. Glanz F. M. Lewis & G. K. Rimer. (Eds.), *Health Behavior and Health Education* (pp. 39–62). San Francisco, CA: Jossey-Bass.
- SAS (Version 9.1) [Computer Program]. (2005). Cary, NC: SAS Institute Inc.
- Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square test statistics in covariance structure analysis. *ASA 1988 Proceedings of the Business and Economic Statistics Section*, Alexandria, VA: American Statistical Association, pp. 308–313.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye, & C. C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422–445.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology 1982* (pp. 290–312). Washington, DC: American Sociological Association.
- Steiger, J. H., & Lind, J. C. (1980, May). Statistically based tests for the number of factors. Paper presented at the annual spring meeting of the Psychometric Society, Iowa City, IA.
- Wayment, H. A., Wyatt, G. E., Tucker, M. B., Romero, G. J., Carmona, J. V., & Newcomb, M., et al. (2003). Predictors of risk and precautionary sexual behaviors among single and married white women. *Journal of Applied Social Psychology*, *33*, 791–816.
- Winship, C., & Mare, R. D. (1983). Structural equations and path analysis with discrete data. *American Journal of Sociology*, *89*, 54–110.
- West, S. G., & Aiken, L. S. (1997). Towards understanding individual effects in multiple component prevention programs: Design, analysis strategies. In K. Bryant, M. Windle, & S. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research*. (pp. 167–209). Washington, D. C.: American Psychological Association.
- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. (pp. 56–75). Thousand Oaks, CA: Sage Publications.
- Xie, Y. (1989). Structural equation models for ordinal variables: An analysis of occupational desination. *Sociological Methods and Research*, *17*, 325–352.
- Yuan, K. -H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology 2000*, *30*, 167–202.