

Archives of Scientific Psychology

http://www.apa.org/pubs/journals/arc



Reporting Results From Structural Equation Modeling Analyses in *Archives of Scientific Psychology*

Rick H. Hoyle and Jennifer C. Isherwood Duke University

ABSTRACT

Psychological research typically involves the analysis of data (e.g., questionnaire responses, records of behavior) using statistical methods. The description of how those methods are used and the results they produce is a key component of scholarly publications. Despite their importance, these descriptions are not always complete and clear. To ensure the completeness and clarity of these descriptions, the *Archives of Scientific Psychology* requires that authors of manuscripts to be considered for publication adhere to a set of publication standards. Although the current standards cover most of the statistical methods commonly used in psychological research, they do not cover them all. In this article, we propose adjustments to the current standards and the addition of additional standards for a statistical method not adequately covered in the current standards—structural equation modeling (SEM). Adherence to the standards we propose would ensure that scholarly publications that report results of data analyzed using SEM are complete and clear.

SCIENTIFIC ABSTRACT

For the scientific community, we recommend reporting standards consistent with the Journal Article Reporting Standards (JARS) of the American Psychological Association for manuscripts in which results from SEM analyses are presented. For all sections of the general JARS except the results section, we recommend minor adjustments. For the results section of the JARS, we provide a supplemental module specific to reports of research that use SEM. The result is a questionnaire that ensures thorough and detailed reports of SEM analyses in the *Archives*.

Keywords: reporting standards, structural equation modeling, JARS

The American Psychological Association's Publication and Communications Board Working Group on Journal Article Reporting Standards (JARS Group) proposed a set of standards for reports of research involving new data collection or quantitative syntheses of extant research (APA Publication and Communications Board Working Group on Journal Article Reporting Standards, 2008). These standards have been translated into a questionnaire that is completed by the authors of manuscripts submitted to the Archives of Scientific Psychology as a means of ensuring that the reporting standards have been met. By encouraging thoroughness, consistency, and transparency in research reports, the JARS contribute to the quality and potential impact of research by psychological scientists.

Psychological science takes many forms and, for this reason, the JARS Group presented their recommendations as "a work in progress," assuming that new methods and paradigms would emerge as the discipline matures. They also acknowledged that, although most of the JARS are relevant for any data-based research report, it would be necessary to develop supplemental modules for specific research designs and statistical approaches. In this article, we propose minor adjustments to sections of the general JARS questionnaire and present a new supplemental module for reports of research in which structural equation modeling (SEM) is the statistical method by which evidence is brought to bear on a set of hypotheses or research questions.

This article was published July 22, 2013.

Rick H. Hoyle and Jennifer C. Isherwood, Department of Psychology and Neuroscience, Duke University.

For further discussion on this topic, please visit the Archives of Scientific Psychology online public forum at http://arcblog.apa.org.

The first author's effort on this article was supported in part by Grant P30 DA023026 from the National Institute on Drug Abuse (NIDA). The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of NIDA or the National Institutes of Health.

Correspondence concerning this article should be addressed to Rick H. Hoyle, Department of Psychology and Neuroscience, 417 Chapel Drive, Duke University, Durham, NC 27708-0086. E-mail: rhoyle@duke.edu

Why a JARS Module Specific to SEM?

Despite the increasing complexity of psychological theories, accessibility of precise and sophisticated data collection methods, and development of flexible and general statistical models for analyzing data, it remains the case that most psychological studies result in a comparison of mean values on a dependent variable for different groups, points in time, or both. Whether by null hypothesis statistical testing or effect size estimation, a conclusion is drawn regarding whether there is evidence that a single parameter—the difference between means—takes on a value suggesting an effect of statistical or practical significance. For studies of this type, and other types of studies in which the focus is estimating and testing individual parameters (e.g., prediction studies, in which the focus is regression coefficients), the questionnaire based on the reporting standards developed by the JARS Group is appropriate and adequate.

Although SEM could be used solely for estimating and testing individual parameters, it is used to fullest benefit when parameters are estimated and tested in the context of a model fit to data. Modeling is inherently multivariate and, when applied to psychological data, focuses on the degree to which patterns of means and covariation between variables observed in data can be accounted for by statistical models that mirror conceptual models of interest (Rodgers, 2010). Such models, like the conceptual models they reflect, may include many variables that are interrelated in complex ways. The degree to which they account for patterns in observed data typically is evaluated in an absolute sense—Does the model fit the data?—and in a relative sense—Is the fit of the model of interest superior to the fit of rival models? The outcome of model fitting using SEM is a potentially large number of fit statistics, parameter estimates, and standard errors simultaneously estimated by methods for which a consideration of assumptions is critical. The amount of information that could be reported is compounded when multiple models are considered or a single model is considered for different populations. The kind of information and the amount of information expected in reports of modeling using SEM differ significantly from what is expected in the standard research report in psychological science.

Because of the challenges authors face when reporting results of studies in which SEM is used to fit models to data, recommendations and guidelines have been offered in a number of publications. Among the earliest publications of this sort is a brief invited article in Psychology and Aging "in response to confusion regarding ideal methods for reporting structural equation modeling" (Editor's Note in Raykov, Tomer, & Nesselroade, 1991, p. 499). The article offers a set of publication guidelines, ranging from the presentation of models to be fit to evaluation of fit and parameter estimates. A fuller and more detailed set of recommendations is offered in a widely cited chapter aimed at helping researchers pen thorough and complete reports of findings from application of SEM (Hoyle & Panter, 1995). The focus of these publications was primarily the reporting of results. Subsequent publications offered suggestions for interpreting and discussing SEM results (e.g., Boomsma, 2000; Hoyle, 2011). Additional publications targeted researchers in particular disciplines, ranging from psychology (McDonald & Ho, 2002) to education (Schreiber, Nora, Stage, Barlow, & King, 2006) to information systems (Gefen, Rigdon, & Straub, 2011). More recent publications have embedded recommendations for writing about SEM results in broader treatments of best practices (e.g., Mueller & Hancock, 2008); offered suggestions for reviewers of manuscripts reporting SEM results (implying recommendations for manuscript authors, Mueller & Hancock, 2010); and offered suggestions for reporting the results of advanced applications of SEM, which involve considerations beyond those typical of the more basic models assumed by most publications offering recommendations and guidelines (Boomsma, Hoyle, & Panter, 2012).

Although there is a general consistency in the recommendations offered by these authors, those recommendations have not been translated into a set of reporting standards. In the remainder of this article, we propose SEM reporting standards consistent with those recommended by the JARS Group. Specifically, we prescribe minor adjustments to several sections of the general JARS questionnaire and provide a supplement to replace the results section of that questionnaire when results from SEM analyses are reported. The standards cover all important reporting considerations for the most frequent applications of SEM but do not address considerations unique to specialized applications and emerging directions (e.g., latent growth curve models, mixture models, Bayesian SEM), each of which will require a set of supplemental standards as its use becomes more widespread.

Minor Adjustments to Sections of the JARS Questionnaire

Outside of questions relevant to the presentation of results, most of the questions in the general JARS questionnaire are appropriate for reports of research involving SEM. Yet, because the development and fitting of models involves considerations beyond those associated with developing and testing hypotheses about parameters, minor adjustments to sections of the JARS questionnaire other than the results reporting section are warranted. The adjustments we recommend are presented in Table 1.

In most cases, these recommendations do not require elaboration. The most significant recommended adjustment is the addition of a set of questions in the method section of the questionnaire for manuscripts reporting findings from simulation research using SEM. Such studies are published with some frequency in psychological journals and often are highly cited. For example, Bentler's (1990) Psychological Bulletin article on the performance of different fit indexes as a function of sample size boasts nearly 6,000 citations. Similar studies are routinely reported in Psychological Methods and, in some cases, would be appropriate for the Archives. Information about the number of samples, the means by which they were generated, and their suitability for later analysis replaces information about sampling procedures in the standard research report. Apart from these questions for simulation research using SEM, which are suitable for other types of simulation research, the recommended additions and changes to the questionnaire sections other than the results section are minor and relatively straightforward.

A JARS Supplemental SEM Module

As established earlier, the results of modeling by using SEM differ from the results of traditional hypothesis testing both in quality and quantity. As such, we recommend that reports of research for which SEM is used for modeling data forego the results section of the general JARS questionnaire, instead completing the supplemental module displayed in Table 2.

The JARS SEM module is divided into six sections. These provide authors with a means of addressing the major concerns associated with their approach to model fitting; the appropriateness of their data for specific estimators and tests; and the reporting of statistical results. The data preparation section focuses on two important concerns in applications of SEM: (a) Were there missing data and, if so, how were they addressed? (b) Was multivariate

Table 1
Adjustments to Sections Other Than the Results Section of the JARS Questionnaire for All Studies
Reporting Results of Structural Equation Modeling

| TITLE | The multivariate data to which SEM analyses often are applied and the complexity of relations SEM can be used to test makes it unlikely that, in most cases, the variables under investigation and the relations between them could be fully identified in the title. Instead, we recommend identifying the mechanism or process reflected in the primary model to which the data are fit. | | | |
|---------------------|--|--|--|--|
| AUTHOR NOTE | Same as in general questionnaire | | | |
| SCIENTIFIC ABSTRACT | Same as in general questionnaire except replace effect sizes, confidence intervals, and/or significance levels under the findings, including: bullet with two or more omnibus fit indices from different classes? yes no, explain | | | |
| | whether the interpreted model is the originally specified model? yes no, explain | | | |
| | | | | |
| INTRODUCTION | Same as in general questionnaire but replace material under the describe the specific hypothesis or objectives bullet with describe the primary model to be fit to data Theories or other means to derive the model and important paths in it? In manuscript supplemental files not relevant primary paths? In manuscript supplemental files not relevant secondary paths? In manuscript supplemental files not relevant planned re-specification if initially specified model does not fit? In manuscript supplemental files not relevant explain | | | |
| METHOD | Same as in general questionnaire except as follows: In the Sampling procedures section, begin with Were the data collected from research participants or generated by simulation? research participants simulation If the data were collected from research participants, follow the general questionnaire but add to the Measures and covariates section: Were indicators of latent variables drawn from one or more multi-item questionnaires? yes no there are no latent variables in the model If yes, then for each questionnaire Were indicators items or parcels? If parcels, how were they constructed? If items, were they treated as categorical or continuous? | | | |
| | | | | |
| | Replace the How was sample size determined bullet in the Sample size, power, and precision section with the following: On what basis was the target sample size determined? rule of thumb, explain availability or resource constraints, explain power analysis precision of parameter estimate(s) or indices of fit, explain | | | |
| | If determination was based on power analysis, on what hypothesis test did the analysis focus? Explain and justify. omnibus model fit significance of key parameters | | | |
| | what minimum value of power was used? | | | |

If the data were generated by simulation, skip the Sample size, power, and precision; and Measures and covariates sections. Replace the Sampling procedures bullets with the following:

(table continues)

Table 1 (continued)

| | What software/algorithm was used for simulation? |
|------------|--|
| | |
| | • State and justify the number and size(s) of samples generated. |
| | |
| | Were any samples lost due to nonconvergence or inadmissible estimates? yes, explainno |
| | |
| DISCUSSION | Same as in general questionnaire, except add to the list under the Are results interpreted taking into account bulle number of and basis for model modifications? In manuscript supplemental files not relevant equivalent/alternative models? In manuscript supplemental files not relevant explain |
| | |

normality evaluated and, if the multivariate distribution was not normal, how was this addressed? Both of these concerns could be addressed either at the stage of data preparation (e.g., multiple imputation and variable transformations, respectively) or model estimation (e.g., full-information maximum likelihood estimation or distribution-free estimation, respectively). Regardless of how these concerns, if present, are addressed, they are evaluated prior to model fitting and, therefore, are appropriately covered in the report prior to the presentation of results. (See Malone & Lubansky, 2012, for specific evaluation guidelines and strategies.)

Because of the many types of models that can be fitted using SEM and the implications of different approaches to specification for estimation and fit, reports of results from SEM analyses should provide ample detail regarding model specification. The principal consideration is whether the model specification is described in sufficient detail that a reader could specify the model and, using the authors' data, obtain the results presented in the report. If, for example, the model includes latent interaction effects, given the multiple ways in which such effects could be specified, a description that simply notes that the model included one or more latent interaction effects is not sufficient. Details regarding the specification strategy used by the author of the report should be provided. Additional information of import for evaluating an application of SEM includes detail regarding the model to which the data are fitted (e.g., degrees of freedom, identification) and justification for aspects of the model that constitute hypotheses (e.g., assignment of indicators to latent variables, covariances between uniquenesses). Hoyle (2012) provided details about these and other specificationrelated matters.

As shown in Table 2, the remaining concerns to be addressed in the research report prior to the presentation of statistical results are estimation, evaluation of fit, and, if relevant, respecification. For estimation and evaluation of fit, the primary concerns are documentation and justification of choices made between available alternatives (e.g., maximum likelihood vs. weighted least squares). For respecification, the primary concern is full-disclosure regard-

ing the basis for model modifications and conceptual justification for modifications. Ample information regarding the choices, decision criteria, and typical justifications are available (e.g., Chou & Huh, 2012; Lei & Wu, 2012; West, Taylor, & Wu, 2012).

The heart of the results section is, of course, the presentation of statistical results. This section of the SEM JARS supplement is the longest and most prescriptive. It is designed to ensure that authors provide adequate detail regarding the fit of the model as a whole (including comparison with rival and respecified models) and the estimates and tests of individual parameters in the model. On the latter count, we recommend reporting of all estimated parameters, not just those of direct relevance to the conceptual focus of the research. Such reporting allows the astute reader to detect any problems with the data or specification that might manifest only as inadmissible estimates or suspect standard errors. A detailed example showing how results for a relatively complex model can be presented in compact form is provided by Hoyle (2011, Chapter 4).

Summary and Conclusions

The JARS questionnaire completed by authors submitting manuscripts to the Archives ensures that readers are provided adequate information for judging the quality of the ideas and findings presented in the manuscript. Although the general JARS questionnaire is adequate for many, perhaps most, reports of research in psychological science, it does include queries relevant for certain specialized methods. The standards for adequate reporting of findings from SEM analyses share a great deal in common with the standards for all reports of psychological science; however, SEM analyses involve important decisions and choices that are unique. For that reason, we developed and presented in this article additional standards for reports of findings from studies that use SEM. We recommend minor adjustments to all sections of the JARS questionnaire except the results section, which we recommend replacing with the supplemental module described in this article. Inclusion of these adjustments and use of the SEM-specific module

Table 2

HOYLE AND ISHERWOOD

| RESULTS | Please provide the information requested in this supplemental JARS questionnaire, or in the text box provide the page number, table, or supplemental file in which the information can be found. | | | | |
|------------------|--|--|--|--|--|
| Data Preparation | If some data are missing, | | | | |
| | state the percentage of missingness and how it is distributed across cases and variables. | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | is it plausible that the data are at least missing at random? yes, explain no | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | o indicate which method was used to address missingness. | | | | |
| | multiple imputation FIML substitution of values deletion of cases other (describe below) | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | Were the data evaluated for multivariate normality? yes no, explain | | | | |
| | yes no_, explain | | | | |
| | | | | | |
| | | | | | |
| | If data are not multivariate normal, what strategy was used to address nonnormality? | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| Specification | Which general approach best describes the use of SEM? strictly confirmatory alternative models model-generating | | | | |
| | Is a path diagram provided for each model fit to the data? | | | | |
| | yes no, explain | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | Is a full account of the model specification for all models to be evaluated provided, including: latent variables? yes no | | | | |
| | fixed and free parameters? yes no constrained parameters? yes no | | | | |
| | | | | | |
| | Is sufficient information provided that, for all models evaluated, the degrees of freedom can be derived by the reader? | | | | |
| | yes no | | | | |
| | Is model identification addressed? yes, but not established yes, and established no , explain | | | | |
| | yes, our not communicate. Yes, and communicate in the separation | | | | |
| | | | | | |
| | | | | | |
| | Was the method for establishing model identification stated and justified? | | | | |
| | yes no, explain not established | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

yes__ no__ no measurement component__

(table continues)

Table 2 (continued)

| | • | If the model includes a means component, is the specification of the mean structure described fully? yes no no means component | | |
|-------------------|---|--|--|--|
| | | | | |
| | • | If the model includes interaction effects, explain how those effects were specified. interaction effects no interaction effects | | |
| | | | | |
| | • | If the data are nested (e.g., occasion within person, student within classroom), explain how nonindependence is accounted for in the model. nested data no nested data | | |
| | | | | |
| | • | Are any comparisons of parameters to be made between groups or occasions? If so indicate which parameters are to be compared for which groups or occasions. yes no | | |
| | | | | |
| Estimation | • | Was the software (including version) used for estimation noted? yes no | | |
| | | | | |
| | • | Which estimation method was used? Justify its use. | | |
| | | | | |
| | • | Were any default criteria (e.g., number of iterations, tolerance) adjusted in order to achieve convergence? yes, explain no | | |
| | | | | |
| | • | Is there any evidence of an improper solution (e.g., error variances constrained at zero; standardized factor loadings greater than 1.0)? yes, explain no | | |
| | | | | |
| Evaluation of Fit | • | How was omnibus fit evaluated (statistics/indexes and criteria)? | | |
| | | | | |
| | | Were criteria clearly stated and adhered to for all evaluations of fit? yes no, explain | | |
| | | | | |
| | • | If alternative models were compared, what strategy and criterion was used to select one over the other(s)? alternative models no alternative models | | |
| | | | | |
| | • | If individual parameters were tested, what test and criterion was used? | | |

(table continues)

HOYLE AND ISHERWOOD

Table 2 (continued)

| includii | neters were compared between groups or occasions, indicate how those comparisons were made, ng criterion. n groups/occasions comparisons no comparisons |
|-----------|--|
| | |
| yes | |
| 0 | |
| 0 | state clearly which fixed parameters were freed or fixed in order to produce the interpreted model(s) |
| | |
| 0 | provide a theoretical or conceptual justification for parameters that were freed or fixed after specification searching |
| | ollowing information provided in the manuscript? |
| 0 | covariance matrix, or a correlation matrix with standard deviations yes no, explain |
| O | means yes no, explain |
| 0 | univariate skewness and kurtosis values yes no_, explain |
| | |
| readers | case-level data archived and information provided so that they could be accessed by interested ? no_, explain |
| • Does th | the manuscript clearly indicate whether the model(s) for which results are presented were specified |
| before of | no, explain |
| For all | models to be interpreted, does the report provide |
| 0 | |
| | (table continu |
| | Was on yes_ If yes, t Are the readers yes_ Does the before yes_ For all |

Table 2 (continued)

| Presentation of Results (continued) | O | difference tests for comparisons between alternative models? yes no _, explain |
|--|----------------------|---|
| | | |
| | | |
| | | |
| | • For all est | imated parameters, does the report provide estimates, indicating whether they are unstandardized or standardized? |
| | 0 | yes_ no_, explain |
| | | |
| | ٥ | standard errors with unstandardized estimates or, with standardized estimates, results of statistical tests? yes no, explain |
| | | |
| | ٥ | cutoffs for standard levels of significance? yes no, explain |
| | | |
| | yes no | |
| | If yes, doe | s the report provide parameter estimates? |
| | | yes no_, explain |
| | | |
| | 0 | state and justify the chosen strategy for testing the effect? yes no_, explain |
| | | |
| | yes no If yes, do | significant interaction effects? Tollow-up analyses make clear the underlying pattern?, explain |
| | | |
| | | |

will ensure that reports of results from SEM analyses in the *Archives* are consistent in style and thoroughness with the recommendations of the JARS Group.

References

- APA Publication and Communications Board Working Group on Journal Article Reporting Standards. (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist*, 63, 839–851.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238–246.
- Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling*, 7, 461–483.
- Boomsma, A., Hoyle, R. H., & Panter, A. T. (2012). The structural equation modeling research report. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 341–358). New York, NY: Guilford Press.
- Chou, C.-P., & Huh, J. (2012). Model modification in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 232–246). New York, NY: Guilford Press.
- Gefen, D., Rigdon, E. E., & Straub, D. (2011). An update and extension to SEM guidelines for administrative and social science research. *Management Information Systems Quarterly*, 35, A1–A7.

- Hoyle, R. H. (2011). Structural equation modeling for social and personality psychology. London, England: Sage.
- Hoyle, R. H. (2012). Model specification in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 126–144). New York, NY: Guilford Press.
- Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 158–176). Thousand Oaks, CA: Sage.
- Lei, P-W., & Wu, Q. (2012). Estimation in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 164–179). New York, NY: Guilford Press.
- Malone, P. S., & Lubansky, J. B. (2012). Preparing data for structural equation modeling: Doing your homework. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 263–276). New York, NY: Guilford Press
- McDonald, R. P., & Ho, M. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, 7, 64–82.
- Mueller, R. O., & Hancock, G. R. (2008). Best practices in structural equation modeling. In J. W. Osborne (Ed.), Best practices in quantitative methods (pp. 488–508). Thousand Oaks, CA: Sage.
- Mueller, R. O., & Hancock, G. R. (2010). Structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), The reviewer's guide to quantitative methods in the social sciences (pp. 371–383). New York, NY: Routledge.

Raykov, R., Tomer, A., & Nesselroade, J. R. (1991). Reporting structural equation modeling results in Psychology and Aging: Some proposed guidelines. *Psychology and Aging*, 6, 499–503.

Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist*, 65, 1–12. Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99, 323–337.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York, NY: Guilford Press.

Received August 8, 2012
Revision received June 7, 2013
Accepted June 8, 2013