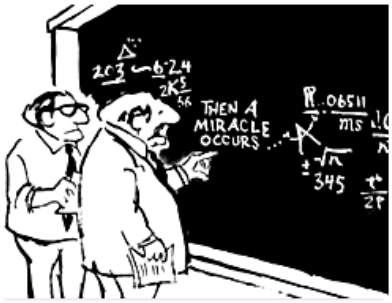
Advanced Data Analysis: Comm 597B

Fall 2014

Dr. Mary Beth Oliver
210 Carnegie Building
571-6830 (C)
mbo@psu.edu



Office Hours:
Tuesdays: 12-1; 4:30-5:30
(sign up)
And by Appointment

Purpose of Course:

Structural equation modeling (SEM) and related procedures have become very popular techniques in the field of Communication and similar social scientific disciplines—and for good reason! These techniques can handle almost any and all analytic approaches that are commonly employed (e.g., regression, factor analysis, analysis of variance), but they can do so much more. SEM and related procedures allow scholars to model and test theoretical relationships between variables, they acknowledge and account for the fact that we routinely work with unobserved variables that are measured with error, and they force us to "see" and "think" in theoretical terms. As a result, a basic understanding of SEM has now become required and assumed among Communication scholars employing quantitative methodologies—it's routinely reported in journals, it's often requested by reviewers, and it opens up more theoretically enriching examinations of our data.

Although SEM is an evolving, vast, and hotly debated area, the basic processes involved are ones that are generally familiar and accessible to scholars and students with solid statistical backgrounds. Consequently, the purpose of this course is to provide you with an introduction to and foundation for SEM that is contextualized in terms of applied research that mirrors the types of questions routinely encountered by Communication scholars. To those ends, this course will emphasize a conceptual understanding (rather than a mathematically derived focus) of the processes involved and decisions required in conducting these types of analyses, it will illustrate how researchers often report their results in scholarly publications, and it will provide you numerous opportunities to practice your skills, both during class and on your own.

It is important that students in the class recognize that this course is introductory and applied. There are many, many topics in SEM that we will not have time cover. It is my hope that this course will provide you with a solid foundation to explore these additional topics either on your own or in more formal study in the future.

Prerequisites:

At present there are no formal prerequisites for the course. However, this course assumes that you have some prior statistical training. If you don't have any background in statistics, you should take a more basic course prior to taking this one. This class employs SPSS (using syntax). If you are completely unfamiliar with SPSS, you will find it much easier to follow along if you spend some time familiarizing yourself with the SPSS interface at the very start of the semester.

Required Texts:

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.
Journal Articles to Be Announced

Evaluation:

Exam 1: In Class: 30%
Exam 2: In Class: 30%
Final Group Project: 30%
Practice Problems/Responses: 10%

Components of Evaluation:

Exams:

Two take-home exams will be administered during the course of the semester. These exams will be open-book, open-note, but they will completed individually. These exams will ask you to apply the concepts that you have learned in the course, to justify your decisions in analysis, to present your data, and to interpret others' research.

- Although the exams are open-book and open-note, it is imperative that you prepare for the exams by reviewing course exercises, homework assignments, and your readings.
- These exams are cumulative.

Final Group Project:

The final group project (in groups of 3-5 people/group) will involve an analysis, write-up, and presentation of some hypotheses/research questions that employ your own data (or archival data).

The paper itself will consist of the following elements:

1) A 1-page rationale for the research, followed by your hypotheses/research questions.
2) A description of the analytic strategy employed (probably about 1-2 paragraphs).
3) A results section (probably about 2-4 pages).
4) A 1-page interpretation of your findings.
5) References (you will need to include references in APA style).

The paper ***must*** be written in APA style and take the format of a journal-article submission. One goal of the project is for you to produce a document that may be feasibly submitted to a journal/conference.

You can use your own data for the project *or* you can use existing (archival or secondary) data. However, please note that if you analyze existing data, *you still need to get IRB approval*. That is, you cannot publish or submit your work to a conference without IRB approval if you analyze archival (secondary) data. In general (though not always), this type of research is considered "exempt."

The presentation of the research will involve a 10-minute overview of your work (probably presented on PowerPoint) during the last day of class. This presentation should take the format of a conference presentation. Following the presentation, we will engage in a brief Q&A session. You should be prepared to clarify your research, entertain questions, and elaborate on any issues that arise during the Q&A.

All materials associated with your project ***must be turned in by the last day of class***. These materials consist of:

- Your paper
- Your data files
- Your data models and/or syntax
- Your presentation

Evaluation will be based on the following criteria:

- Completeness (all required components present)
- Clarity and organization of presentation (including APA style)
- Defensible data-analytic strategy
- Accuracy of analysis and presentation
- Clarity and appropriate format of presentation
- Group Contribution (based on peer review)

"Homework" (not graded):

Because comfort and skill in data analysis increase dramatically with practice, every class session will include "homework" exercises. This homework will require you to apply the skills that you learned in class and from your texts, to interpret results reported by other authors, and to illustrate your decision-making in the data analysis process. These homework exercises will not be graded. However, my recommendation is that you complete the homework as if you were receiving a grade on it, and subsequently check your answers – first with classmates and then with the answers that I provide. Of course, whether or not you choose to do the "homework" is up to you, but working through the problems will help you prepare for the exams and increase your ability and confidence.

Practice Problems:

Although working through the homework problems will provide you with good practice, it is my experience that the act of *creating* practice problems further enhances familiarity and expertise. Consequently, after each class session (beginning with the 2nd week of class) you will be expected to submit your own practice problem (and answers) onto the Angel website OR respond to a classmate's posting (alternate weeks for submissions/responses).

- *On the weeks you're submitting a problem*: Practice problems and answers can be modeled after the homework, or they can be of your own creation. You can work on these with a classmate, but each person must submit a different problem/exercise. I have provided an example on Angel to start you off (which is really the homework for the week). **Practice problems are due by Monday at 5 pm.**
- *On the weeks you're responding*: You should log into Angel and select a post that has not yet received a response. You should indicate that you are posting a class response, and then you should proceed to write your comments. What are you looking for? You should be looking to see if there are errors in the practice

problem, if there are questions that remain, or if there are issues that are unclear. You can also provide positive feedback on a problem that's well done! **Responses are due by Wednesday at noon.**

Aside from the work that is for credit, you should make it a habit to look at all of the practice problems, noting if there are any errors (or thanking them for a good problem!).

Each problem and response that you submit is worth 1 point. I will not be providing comments on these, but will be randomly checking to see that they are being done well. If I see that the problems reflect carelessness or lack of effort, I will let you know so that points won't be deducted. **Late practice problems or responses will never be accepted, so don't wait until the last minute to do these.** Please know that I may use these practice problems in future courses (with due acknowledgement), so don't create practice problems that you feel are proprietary (e.g., that use data that you don't want to share, etc.).

- People with last names beginning A – J:

Do the Practice Problems (Due Mondays by 5 pm)	Respond to Others' Problems (Due Wednesdays by noon)
Week 2 – Review of Basic Stats	Week 3 – More on Regression & Interactions
Week 4 – Regression and Interactions	Week 5 – Intro to Logic of SEM
Week 7 – Path Analysis 1	Week 9 – Path Analysis I
Week 10 – Measurement Models I	Week 11 – Measurement Models II
Week 12 – Structural Equation Models	Week 15 – Multiple Group Analysis

- People with last names beginning K – Z:

Do the Practice Problems (Due Mondays by 5 pm)	Respond to Others' Problems (Due Wednesdays by noon)
Week 3 – More on Regression & Interactions	Week 2 – Review of Basic Stats
Week 5 – Intro to Logic of SEM	Week 4 – Regression and Interactions
Week 9 – Path Analysis I	Week 7 – Path Analysis 1
Week 11 – Measurement Models II	Week 10 – Measurement Models I
Week 15 – Multiple Group Analysis	Week 12 – Structural Equation Models

Course Policies

- It is very important that everyone in the class feel completely free and comfortable to ask questions and contribute to class discussions. I never consider your questions misplaced or misinformed, and you should understand that if you have a question or need clarification, you are most certainly not alone.
- Attendance in every class is expected. Because this course will proceed at a fairly rapid pace, if you miss a class, you miss a lot of information. Even if you are only late for class, you are likely to miss relevant information. Consequently, it is in your interest to be in attendance and to be on time. If you miss a class or are late, please do not ask the instructor to repeat lectures, to go over missed information or procedures, or to accept late work.

- It is expected that you will have read all of the assigned readings prior to coming to class.
- Although I encourage students in the course to study and practice data analysis together, the exams must reflect only your own work. Collaboration on exams will result in a failing grade in the class.
- Students are strongly discouraged from requesting incompletes in this course. Incompletes will be awarded only under extraordinary circumstances (e.g., serious illness or death of a family member). If an incomplete is requested and awarded, the student and I will agree upon an acceptable time frame for the work to be completed. Work that is not completed within this time frame will receive zero points. Please know that incomplete grades ("I's") automatically turn into "F's" after one semester. Once a grade has turned into an "F," I will not complete a change of grade.
- This class will make use of as many "in-class demos and exercises" as possible. Consequently, please make a habit of bringing a calculator with you to class.
- Because we're in a computer lab, I realize that there is a temptation to make use of the facilities in a way unrelated to class (e.g., checking of e-mail). However, I have also noticed that people who engage in this behavior typically slow the class down because they miss important points and then need to ask for clarification. In addition, students sitting around the "offender" are often distracted. Consequently, please do not use the computers for personal reasons during class time. If you do so, I will politely ask you to leave the room.
- No food or drink is allowed in the lab.

Academic Integrity:

Academic integrity is the pursuit of scholarly and creative activity in an open, honest and responsible manner, free from fraud and deception, and is an educational objective of the College of Communications and the University. Cheating – including plagiarism, falsification of research data, using the same assignment for more than one class, turning in someone else's work, or passively allowing others to copy your work – will result in academic penalties at the discretion of the instructor. In serious cases, it could also result in suspension or dismissal from the University or in the assignment of an "XF" grade (failed for academic dishonesty).

As students studying communications, you should understand and avoid plagiarism (presenting the work of others as your own.) A discussion of plagiarism, with examples, can be found at http://tlt.psu.edu/plagiarism/student-tutorial/. The rules and policies regarding academic integrity should be reviewed by every student, and can be found online at: www.psu.edu/ufs/policies/47-00.html#49-20, and in the College of Communications document, "Academic Integrity Policy and Procedures." Any student with a question about academic integrity or plagiarism is strongly encouraged to discuss it with his or her instructor.

Note To Students With Disabilities:

Penn State welcomes students with disabilities into the University's educational programs. If you have a disability-related need for reasonable academic adjustments in this course, contact the Office for Disability Services, ODS located in room 116 Boucke Building at 814-863-1807(V/TTY). For further information regarding ODS, please visit its website at www.equity.psu.edu/ods/. Instructors should be notified as early in the semester as possible regarding the need for reasonable academic adjustments.

Class Schedule

Date		Topic	Kline	Directly Relevant	Data Sources	Additional Resources
Week 1 --	28-Aug	Introduction	Ch. 1			
Week 2 --	4-Sep	Review of Basic Statistics	Chs. 2 & 3			
Week 3 --	11-Sep	More on Regression	Review Ch.			
Week 4 --	18-Sep	Regression and Interactions		Fitzsimons (2008) Hayes & Matthes (2009) Hayes et al. (2012)		
Week 5 --	25-Sep	Intro to the Logic of SEM	Ch. 4	Lei & Wu (2007)		Bentler & Chou (1987) Holbert & Stephenson (2002) Stephenson et al. (2006)
Week 6 --	2-Oct	Exam 1				
Week 7 --	9-Oct	Path Analysis I	Ch. 5		Beullens & VandenBulck (2007)	Hu & Bentler (1999) MacCallum et al. (1993)
Week 8 --	16-Oct	"Work Day"				
Week 9 --	23-Oct	Path Analysis II***	Ch. 6	Hayes (2009)	Beullens &	Baron & Kenny (1986)
Week 10 --	30-Oct	Measurement Models I***	Ch. 7			
Week 11 --	6-Nov	Measurement Models II	Review Ch.	Hayes et al. (2005)	Shen et al. (2008) Wang & Senecal (2008)	Brown (2003) Graham (2006)
Week 12 --	13-Nov	Structural Equation Models	Ch. 8		Lee & LaRose (2007)	Coffmann & MacCallum
Week 13 --	20-Nov	Exam 2				
Week 14 --	27-Nov	Thanksgiving Break				
Week 15 --	4-Dec	Multiple Group Analysis	Ch. 11		Kerkof & Finkenauer (2008)	Marcoulides et al. (2008)
Week 16 --	11-Dec	Group Presentations				

*** These days may have to be modified, either by delivering the lecture online via recorded lectures, and/or by holding the lecture at an alternative time (likely on a Friday afternoon)

References

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality And Social Psychology, 51*, 1173-1182. doi: 10.1037/0022-3514.51.6.1173

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*, 78-117. doi:10.1177/0049124187016001004

Beullens, K., & Van den Bulck, J. (2007, May). *News, music videos, and action-movie exposure and adolescents' intentions to take risks in traffic.* Paper presented at the annual conference of the International Communication Association, San Francisco.

Brown, T. A. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire: Multiple factors or method effects? *Behaviour Research and Therapy, 41*, 1411-1426. doi: 10.1016/S0005-7967(03)00059-7

Coffman, D. L., & MacCallum, R. C. (2005). Using parcels to convert path analysis models into latent variable models. *Multivariate Behavioral Research, 40*, 235-259. doi: 10.1207/s15327906mbr4002_4

Fitzsimons, G. J. (2008). Death to dichotomizing, *35*(1), 5-8. doi: dx.doi.org/10.1086/589561

Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement, 66*, 930-944. doi:10.1177/0013164406288165

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549-576. doi:10.1146/annurev.psych.58.110405.085530

Hayes, A. F. (2009). Beyond Baron and Kenny. *Communication Monographs*, *76,* 408-420. doi: 10.1080/03637750903310360

Hayes, A. F. (2013). *Introduction to moderation, mediation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford.

Hayes, A. F., Glynn, C. J., & Huge, M. E. (2012). Cautions regarding the interpretation of regression coefficients and hypothesis tests in linear models with interactions. *Communication Methods and Measures*, *6*(1), 1-11. doi: 10.1080/19312458.2012.651415

Hayes, A. F., Glynn, C. J., & Shanahan, J. (2005). Willingness to self-censor: A construct and measurement tool for public opinion research. *International Journal of Public Opinion Research, 17*, 298-323. doi: 10.1093/ijpor/edh073

Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods, 41*, 924-936. doi:10.3758/BRM.41.3.924

Hayes, A. F., & Preacher, K. J. (2008). Contemporary approaches to assessing mediation in communication research. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13-54). Thousand Oaks, CA: Sage.

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? , *X*(X), 1-10. doi: 10.1177/0956797613480187

Holbert, R. L., & Stephenson, M. T. (2002). Structural equation modeling in the communication sciences, 1995-2000. *Human Communication Research, 28*, 531-551. doi: 10.1093/hcr/28.4.531

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*, 556-572. doi: 10.1207/s15506878jobem4704_5

Holbert, R. L., & Stephenson, M. T. (2008). Commentary on the uses and misuses of structural equation modeling in communication research. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 185–218). Thousand Oaks, CA: Sage.

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. doi: 10.1080/10705519909540118

Kerkhof, P., & Finkenauer, C. (2008). *Relationship quality and compulsive Internet use: A study among newlywed couples*. Paper presented at the annual conference of the International Communication Association.

Lee, D., & LaRose, R. (2007). A socio-cognitive model of video game usage. *Journal of Broadcasting & Electronic Media, 51*, 632-650. doi: 10.1080/08838150701626511

Lei, P. W., & Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. *Educational Measurement: Issues and Practice, 26*(3), 33-43. doi: 10.1111/j.1745-3992.2007.00099.x

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 151-173. doi: 10.1207/S15328007SEM0902_1

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149. doi: 10.1037/1082-989X.1.2.130

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*, 185-185. doi: 10.1037/0033-2909.114.1.185

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual review of psychology, 58*, 593-614. doi: 10.1146/annurev.psych.58.110405.085542

Marcoulides, G. A., Emrich, C., & Marcoulides, L. D. (2008). Testing for multigroup invariance of the computer anxiety scale. *Educational and Psychological Measurement, 68*, 325-334. doi: 10.1177/0013164407308469

Matsunaga, M. (2008). Item parceling in structural equation modeling: A primer. *Communication Methods and Measures, 2*, 260-293. doi: 10.1080/19312450802458935

O'Keefe, D. J. (2003). Message properties, mediating states, and manipulation checks: Claims, evidence, and data analysis in experimental persuasive message effects research. *Communication Theory, 13*, 251-274. doi: 10.1093/ct/13.3.251

Paek, H.-J. (2008). Mechanisms through which adolescents attend and respond to antismoking media campaigns. *Journal of Communication, 58*, 84-105. doi: 10.1111/j.1460-2466.2007.00375.x

Peter, J., & Valkenburg, P. M. (2008). Adolescents' exposure to sexually explicit internet material and sexual preoccupancy: A three-wave panel study. *Media Psychology, 11*, 207-234. doi: 10.1177/0093650208321754

Reinard, J. C. (2006). *Communication research statistics (Chapter 17: Modeling communication behavior, pp. 451-487)*. Thousand Oaks, CA: Sage.

Shen, L., Condit, C., & Wright, L. (2008, May). *The psychometric property and validation of a fatalism scale*. Paper presented at the annual meeting of the International Communication Association, Montreal, Canada.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422-445. doi: 10.1037//1082-989X.7.4.422

Stephenson, M. T., & Holbert, R. L. (2003). A Monte Carlo simulation of observable versus latent variable structural equation modeling techniques. *Communication Research, 30*, 332-354. doi: 10.1177/0093650203030003004

Stephenson, M. T., Holbert, R. L., & Zimmerman, R. S. (2006). On the use of structural equation modeling in health communication research. *Health Communication, 20*, 159-167. doi: 10.1207/s15327027hc2002_7

Wang, J., & Senecal, S. (2008). Measuring perceived website usability. *Journal of Internet Commerce, 6*(4), 97-112. doi: 10.1080/15533610802052829

Numbers
Mary Cornish

I like the generosity of numbers.
The way, for example,
they are willing to count
anything or anyone:
two pickles, one door to the room,
eight dancers dressed as swans.

I like the domesticity of addition--
add two cups of milk and stir--
the sense of plenty: six plums
on the ground, three more
falling from the tree.

And multiplication's school
of fish times fish,
whose silver bodies breed
beneath the shadow
of a boat.

Even subtraction is never loss,
just addition somewhere else:
five sparrows take away two,
the two in someone else's
garden now.

There's an amplitude to long division,
as it opens Chinese take-out
box by paper box,
inside every folded cookie
a new fortune.

And I never fail to be surprised
by the gift of an odd remainder,
footloose at the end:
forty-seven divided by eleven equals four,
with three remaining.

Three boys beyond their mothers' call,
two Italians off to the sea,
one sock that isn't anywhere you look.