



Structural Equation Modeling for Ecology & Evolutionary Biology

Jarrett E. K. Byrnes
University of Massachusetts Boston
<http://byrneslab.net/teaching/sem>




```

    graph TD
      Lobster[Lobster] --> Urchin[Urchin]
      Sheephead[Sheephead] --> Urchin
      Urchin --> Kelp[Kelp]
      E1[ε₁] --> Urchin
      E2[ε₂] --> Kelp
      Lobster <--> Sheephead
  
```

$$\# \text{ Flowers} = \alpha + \beta * \text{Depth of March Snowpack} + \varepsilon$$


- Describes a system
- But, implies that if we trim flowers, we change March snowpack, right?
- We assume directionality
- There implicit rules underlying =

Mechanistic Description of a System



```

    graph TD
      A[Snowpack on May 15th] -- 0.21 --> B[First Date of Bare Ground]
      B -- 0.54 --> C[Mean Date Of 1st Flowering]
      C -- 0.8 --> D[Total # of Flowers]
      E[# of Inflorescences] -- 11.2 --> D
      A -- 0.6 --> E
      B -.-> D
      C -.-> D
      E1[ε₁] --> B
      E2[ε₂] --> C
      E3[ε₃] --> E
      E4[ε₄] --> D
      D -.-> A
      subgraph N.S.
        B
        C
        D
      end
  
```

Inouye et al 2002 *Oecologia*

Course Goals

1. Give you a working familiarity with SEM
2. Decide when SEM is right for you
3. Understand the process of model creation, evaluation, and revision
4. Be able to implement SEM in R in whatever framework you choose
5. Build models that incorporate real-world complexity to arrive at new insights to scientific challenges

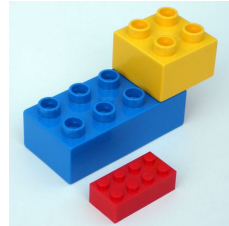
Two Paradigms

Covariance-Based
Estimation (lavaan)

$$\hat{\Sigma} = \begin{Bmatrix} \sigma_{11} \\ \sigma_{12} \sigma_{22} \\ \sigma_{13} \sigma_{23} \sigma_{33} \end{Bmatrix}$$

Implied Covariance Matrix

Local Equation Estimation
(piecewiseSEM)



Testing *piecewiseSEM*

- Use you to test the new version (2.1) of *piecewiseSEM* (which is in development!)



Typical Day

9:00 – 10:30 Lecture/Lab I

10:30 – 11:00 Break!

11:00 – 12:30 Lecture/Lab II

12:30 – 1:30 Lunch

1:30 – 3:00 Lecture/Lab III

3:00 – 3:30 Break!

3:30 – 4:30 Lecture

4:30 – 5:30 Work with Your Data

The Sticky Notes!



- Green/Red for exercises (I know if you're done or are lost)
- Yellow – slow that dangd lecture down!
- Blue – questions for later/at end of day

Thank you, Data Carpentry!

Schedule for the Week

Monday -

What is SEM? A Practical and Historical Overview
SEM Anatomy
Building Multivariate Causal Models
Engines of SEM: Covariance-Based Estimation

Tuesday -

Engines of SEM: Covariance-Based Estimation
What does it mean to evaluate a multivariate hypothesis?
Latent Variable models

...

Schedule for the Week

Wednesday -

Engines of SEM: Local Estimation with piecewiseSEM
A Nonlinear, Non-normal world
Mixed Models in SEM
Categorical Variables and Multigroup Analysis

Thursday -

Spatial Data in SEM
Timseries and Temporal Autocorrelation in SEM
Whole-System Prediction with SEM
Composite Variables

Friday -

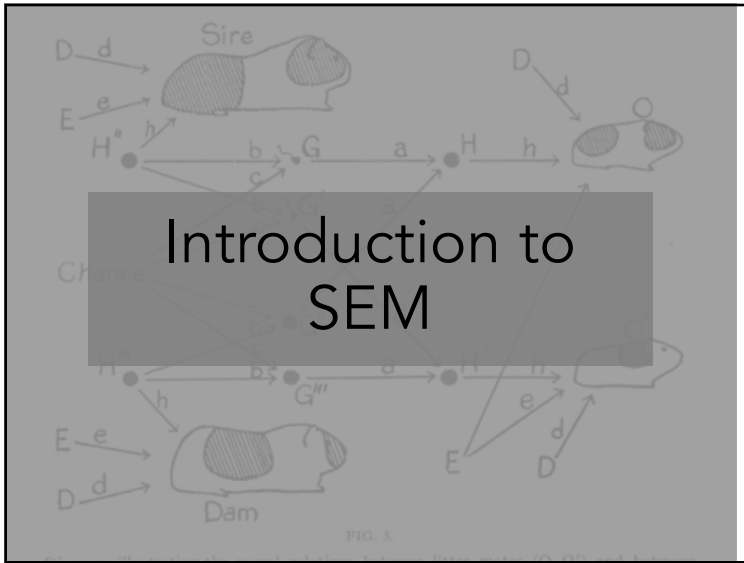
How to Fool Yourself with SEM
Open Lab & Presentations

Who is Jarrett Byrnes?

- Assistant Professor at University of Massachusetts Boston
- Fields:
 - Biodiversity & ecosystem function
 - Kelp forests and global change
 - Salt marsh food webs
 - Ecological statistics



Who are you? Why are you here?



- Overview
1. What is SEM?
 2. A History Lesson
 3. From ANOVA to SEM

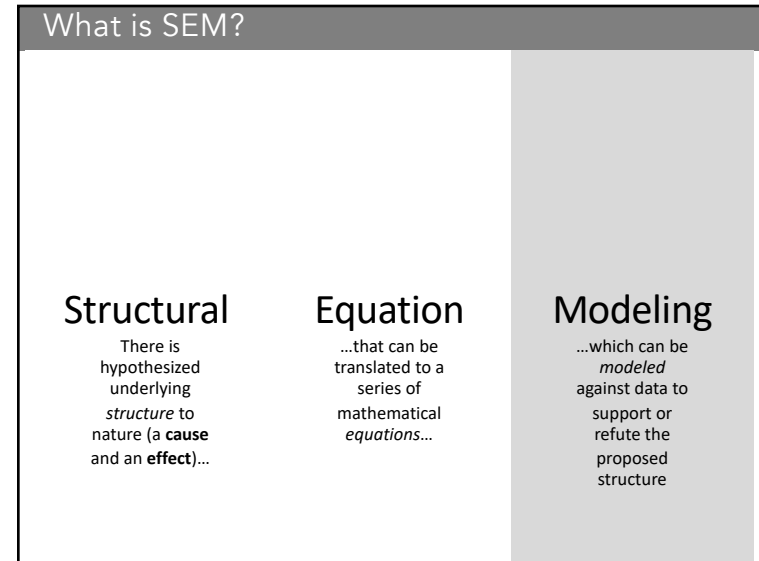
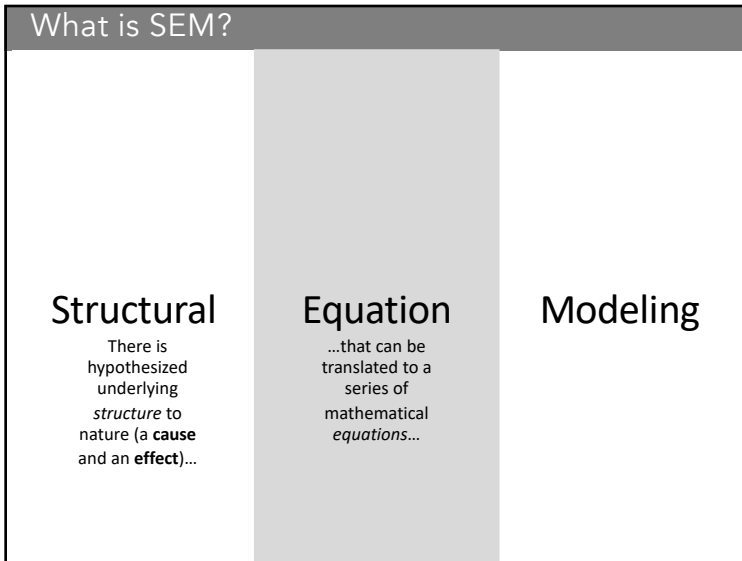
What is SEM?

Structural Equation Modeling

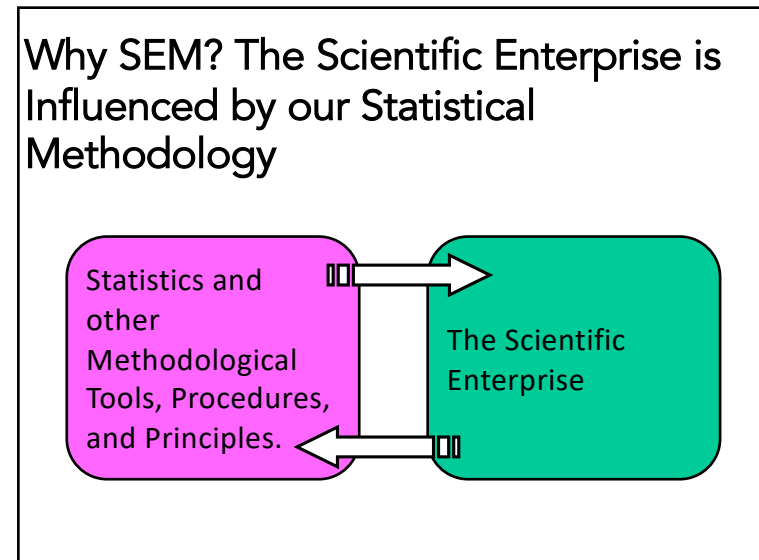
What is SEM?

Structural Equation Modeling

There is hypothesized underlying structure to nature (a **cause** and an **effect**)...



- What is SEM? By any other name ...
- Structural equation modelling (SEM)
 - Path analysis
 - Directed acyclic graph models



What is SEM? A graphical approach

The data

Equation form

$$y_1 = \alpha_{11} + \beta_{11}x_1 + \zeta_1$$

Graphical form

What is SEM? A graphical approach

Charles Elton, "Animal Ecology" (1927, p.58)

FIG. 4.—Food-cycle among the animals on Bear Island, a barren spot in the arctic zone, south of Spitsbergen. (The dotted lines represent probable food relations not yet proved.) The best way to read the diagram is to start at "marine animals" and follow the arrows. (From Summerhayes and Elton.²⁰)

What is SEM? A graphical approach


Paine, 1966

FIG. 1. The feeding relationships by numbers and calories of the *Pisaster* dominated subweb at Mukkaw Bay. *Pisaster*, N = 1049; *Thais*, N = 287. N is the number of food items observed eaten by the predators. The specific composition of each predator's diet is given as a pair of fractions; numbers on the left, calories on the right.

What is SEM? A graphical approach

Wootton (1994) Ecology

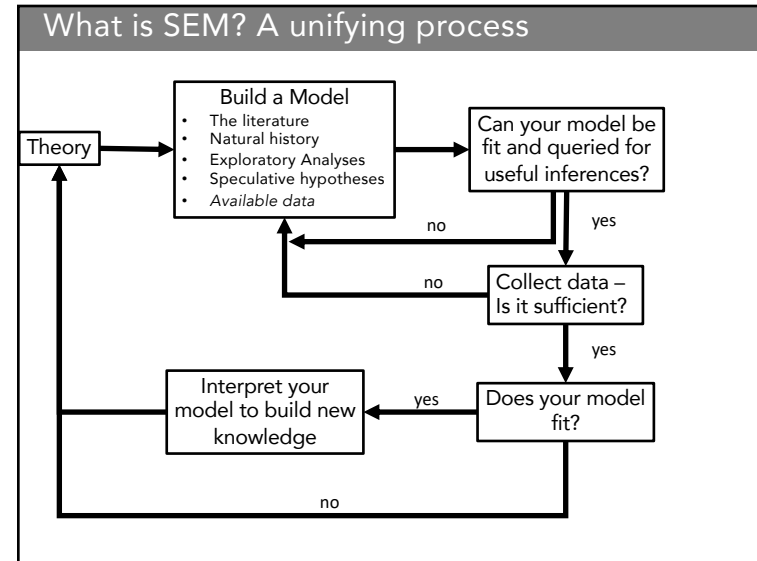
What is SEM? A framework



We use statistical and mathematical tools

within the SEM framework

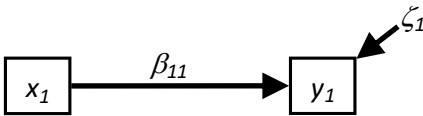
to build scientific understanding about the multiple processes operating in systems



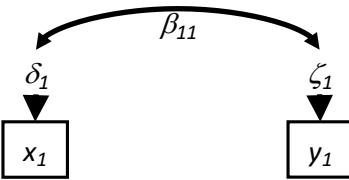
What is SEM? Implies directionality

equation form $y_1 = \beta_{11}x_1 + \zeta_1$

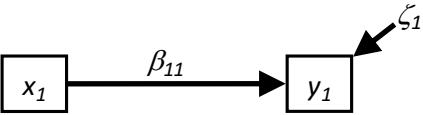
graphical form with causality



graphical form "without" causality



What is SEM? The elephant in the room



"An equation...can be said to be *structural* if there exists sufficient evidence from all available sources to support the interpretation that x_1 has a causal effect on y_1 ."

(Grace, 2006)

What is SEM? What is causation?

- **Key Point #1:** SEM assumes that x causes y
 - Prior observation
 - Prior statistical models
 - Prior experimentation
 - Some or all of the above
- Does not assume ultimate causation

What is SEM? Incorporating complexity

Hypothesis involving indirect effects

Corresponding Equations

e.g.

$$y_1 = \gamma_{11}x_1 + \zeta_1$$

$$y_2 = \beta_{21}y_1 + \gamma_{21}x_1 + \zeta_2$$

$$y_3 = \beta_{32}y_2 + \gamma_{31}x_1 + \zeta_3$$

What is SEM? A complicated network

- **Key Point #2:** By combining inferences across multiple equations, SEM addresses both direct and indirect effects in a system

What is SEM? Indirect effects


What is SEM? Putting it all together

- **Key Point #1:** SEM assumes that x causes y
- **Key Point #2:** By combining inferences across multiple equations, SEM addresses both direct (proximate) and indirect (ultimate) effects in a system


Overview

1. What is SEM?
2. A History Lesson
3. From ANOVA to SEM

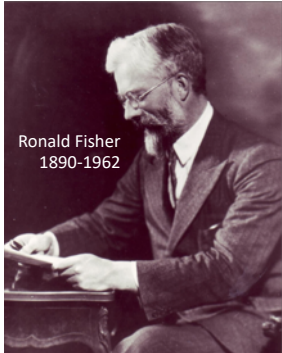
History. Fit, correlation, and testing models



Karl Pearson
1857-1936



Francis Galton
1822-1911



Ronald Fisher
1890-1962

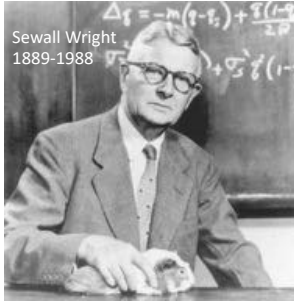
Numerical strength of association (Pearson product moment correlation, r)

Evaluate model fit (Chi-squared goodness of fit, χ^2)

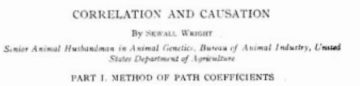
Test hypotheses (ANOVA)

Derive effect sizes (maximum likelihood estimation)

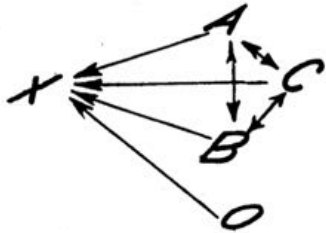
History. Causation vs. correlation



Sewall Wright
1889-1988



CORRELATION AND CAUSATION
By SEWALL WRIGHT
Senior Animal Husbandman in Animal Genetics, Bureau of Animal Industry, United States Department of Agriculture
PART I. METHOD OF PATH COEFFICIENTS



Path analysis

The First Path Diagram?

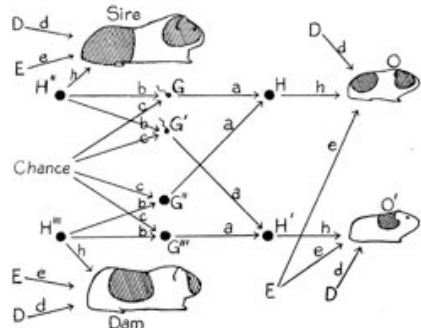


FIG. 3.
Diagram illustrating the causal relations between litter mates (O, O') and between each of them and their parents. H, H', H'', H''' represent the genetic constitutions of the four individuals, G, G', G'', and G''' that of four germ cells. E represents such environmental factors as are common to litter mates. D represents other factors, largely ontogenetic irregularity. The small letters stand for the various path coefficients.

Wright 1920 PNAS

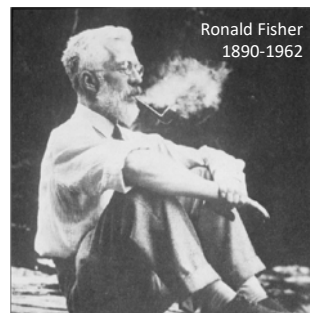
History. Causation vs. correlation

“The basic fallacy of the method appears to be the assumption that it is possible to set up *a priori* a comparatively simple graphic system which will truly represent the lines of action of several variables upon each other, and upon a common result. . . . The pure mathematics by which this is shown is apparently faultless in the sense of algebraic manipulation, but it is based upon assumptions which are wholly without warrant from the standpoint of concrete, phenomenal actuality.” (Niles, 1922)

“The writer has never made the preposterous claim that the theory of path coefficients provides a general formula for the deduction of causal relations. He wishes to submit that the combination of knowledge of correlations with knowledge of causal relations, to obtain certain results, is a different thing from the deduction of causal relations from correlations implied by Niles’s statement. Prior knowledge of the causal relations is assumed as a prerequisite in the former case. Whether such knowledge is ever possible seems to be the subject of Niles’s philosophical discussion of the nature of causation.” (Wright, 1923)

History. Causation vs. correlation

Smoking → Cancer
 Cancer → Smoking
 Smoking ← Gene → Cancer

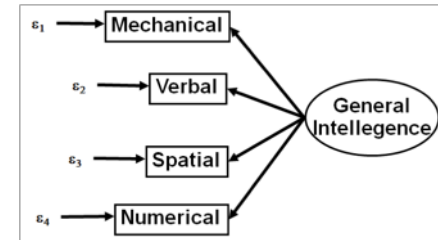


Ronald Fisher
1890-1962

History. Factor Analysis




Charles Spearman
1863-1945

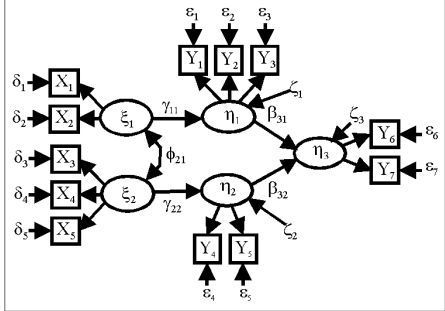


This is the basis of your SAT and GRE score!

History. 2nd Generation SEM


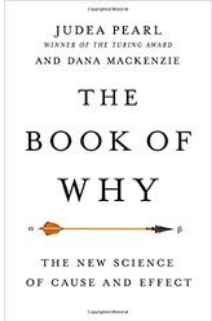
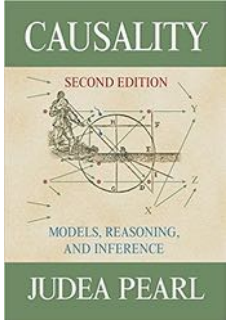


Karl Jöreskog
1935-



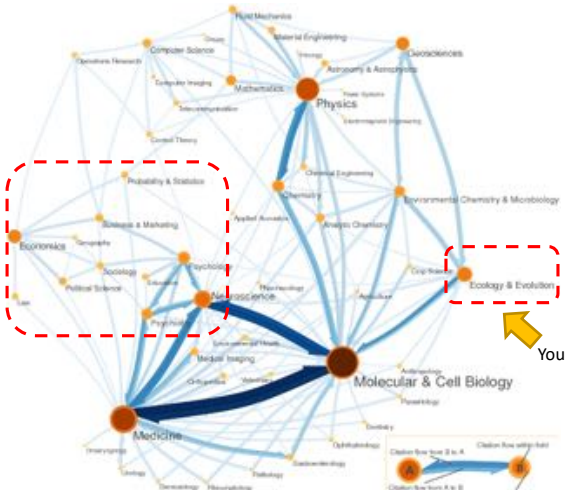
- LISREL = combine path and factor analysis
- Model fit using covariance and ML estimation
- Assess and compare fit of multivariate model

History. 3rd Generation SEM

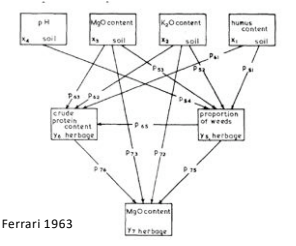
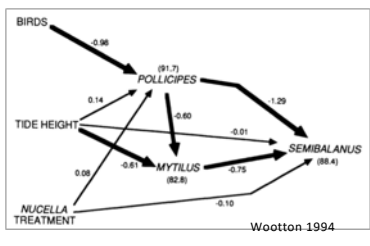
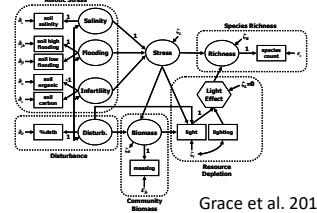
- Unite SEM with graph theory
- Causality is central
- Flexible methods with piecewise approach

History. SEM and Ecology



eigenfactor.org

History. SEM and Ecology

Grace et al. 2010

Methods in Ecology and Evolution

APPLICATION
PIECEWISESEM: Piecewise structural equation modelling in R for ecology, evolution, and systematics

Jonathan S. Lefcheck*
Department of Biological Sciences, Virginia Institute of Marine Science, The College of William & Mary, Gloucester Point, VA 23062-1346, USA

Overview

1. What is SEM?

2. A History Lesson

3. From ANOVA to SEM

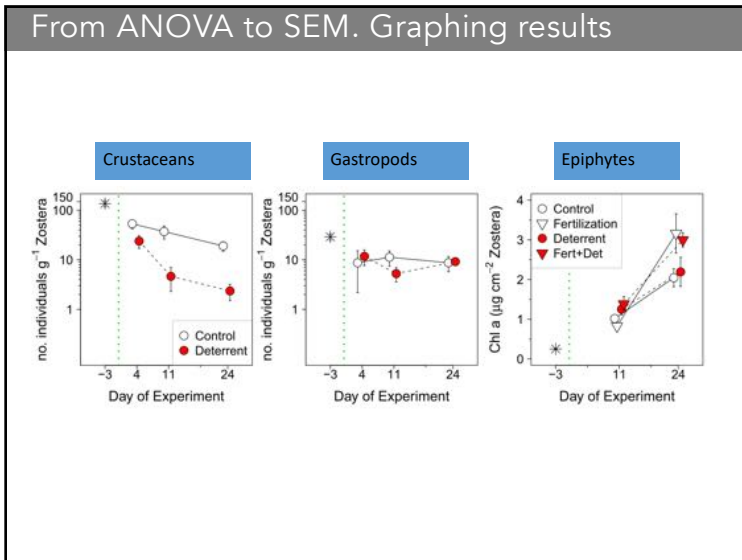
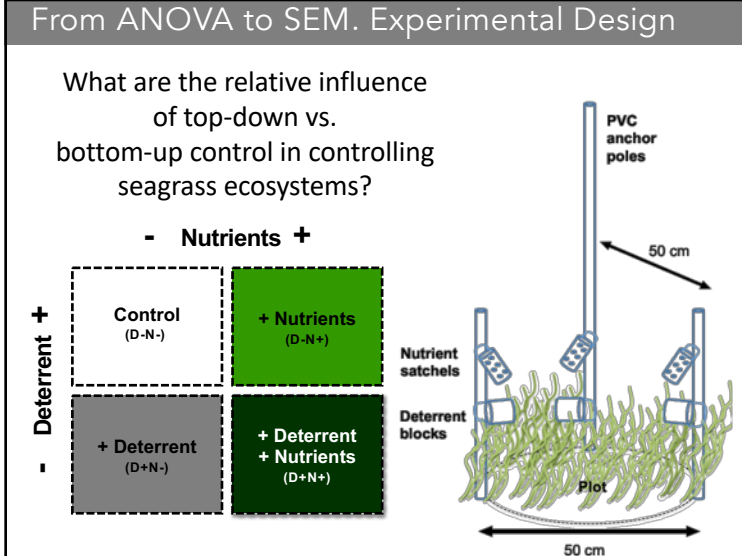
From ANOVA to SEM. Whalen et al. 2013

Ecology, 94(2), 2013, pp. 510-520
© 2013 by the Ecological Society of America

Temporal shifts in top-down vs. bottom-up control of epiphytic algae
in a seagrass ecosystem

MATTHEW A. WHALEN,^{1,3} J. EMMETT DUFFY,¹ AND JAMES B. GRACE²





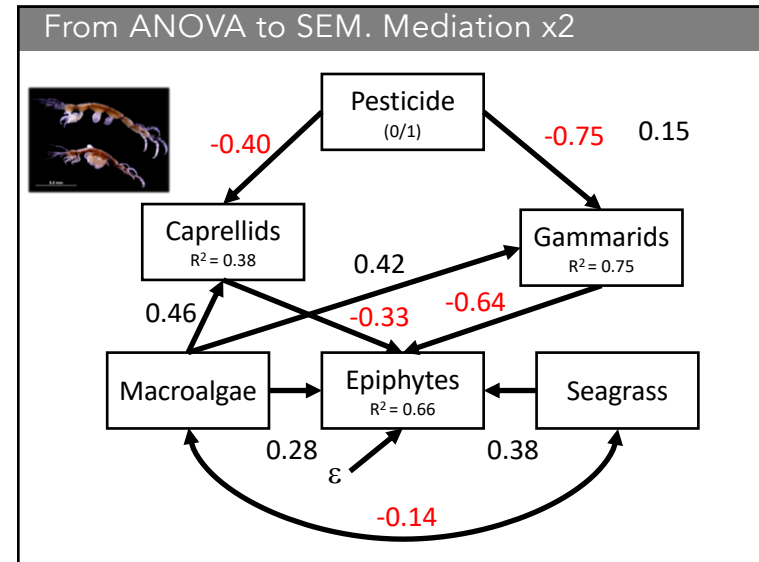
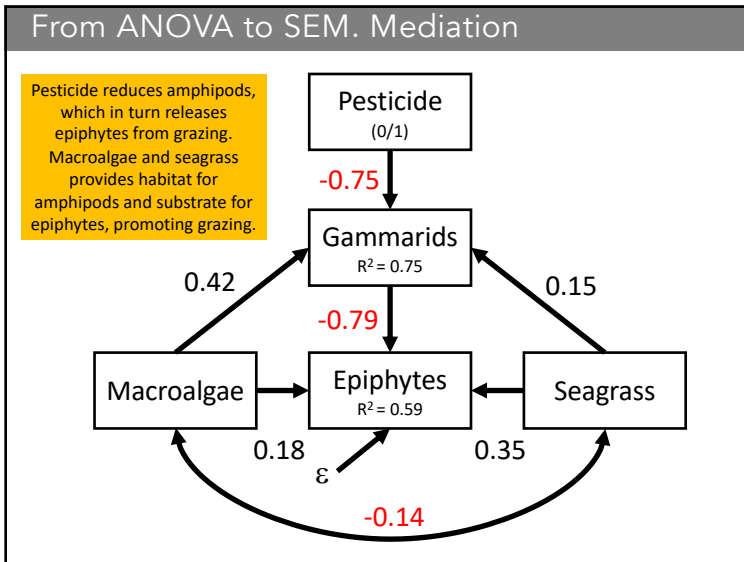
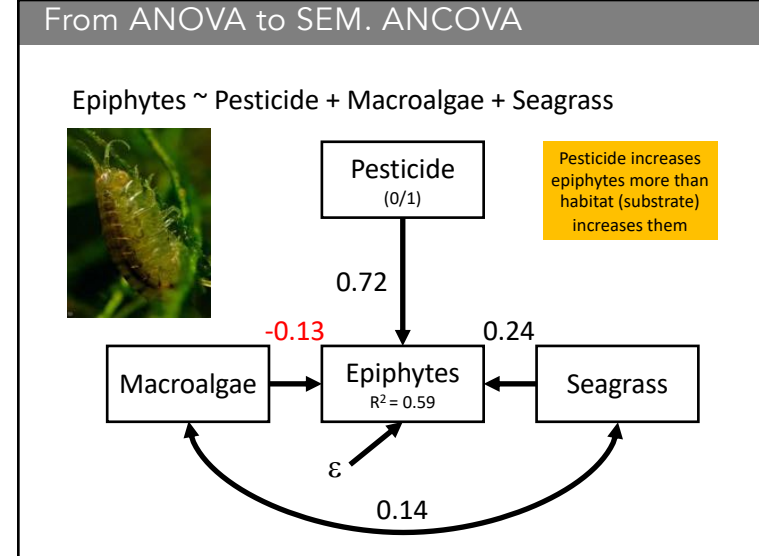
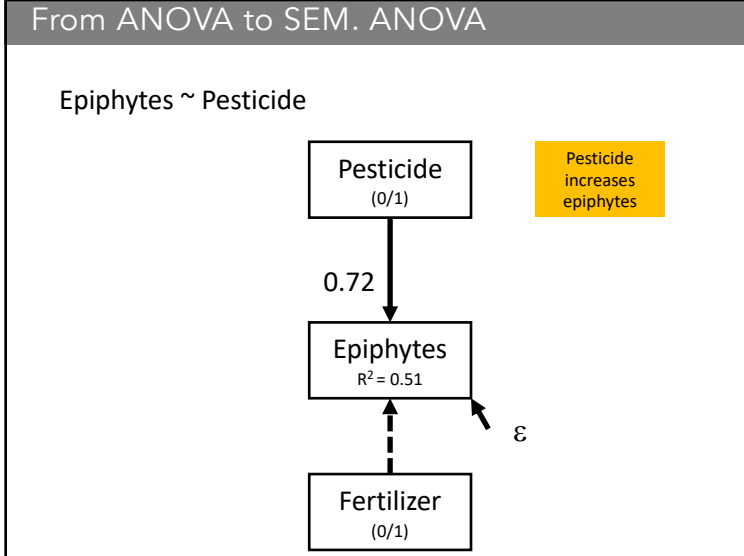
From ANOVA to SEM. Whalen et al. 2013

TABLE 1. Univariate analyses of mesograzer densities and epiphyte biomass from (A) fall and (B) summer experiments in an eelgrass (*Zostera marina*) bed in the York River, Virginia, USA.

Experiment and response source	Crustaceans			Gastropods			Epiphytes		
	df	F	P	df	F	P	df	F	P
A) Fall									
Deterrent	1	42.84	<0.001	1	0.33	0.574	1	3.97	0.052
Fertilization	1	0.00	0.958	1	0.01	0.920	1	3.10	0.084
Sampling date	2	13.77	<0.001	2	0.12	0.887	1	78.24	<0.001
Det. × fert.	1	0.89	0.349	1	11.00	0.002	1	0.86	0.358
Fert. × date	1	0.10	0.756	1	2.00	0.163	1	3.72	0.059
Fert. × date	1	0.58	0.448	1	2.96	0.091	1	7.00	0.011
Det. × fert. × date	1	2.90	0.094	1	0.71	0.403	1	0.81	0.371
Residual	21			51					
B) Summer									
Deterrent	1	129.24	<0.001	1	1.07	0.306	1	66.22	<0.001
Fertilization	1	0.00	0.958	1	0.01	0.920	1	2.19	0.145
Sampling date	1	0.89	0.349	1	11.00	0.002	1	0.83	0.367
Det. × fert.	1	0.10	0.756	1	2.00	0.163	1	1.00	0.322
Det. × date	1	0.58	0.448	1	2.96	0.091	1	6.21	0.016
Fert. × date	1	2.90	0.094	1	0.71	0.403	1	0.53	0.468
Det. × fert. × date	1	1.57	0.216	1	0.27	0.606	1	1.14	0.290
Residual	56			56			56		

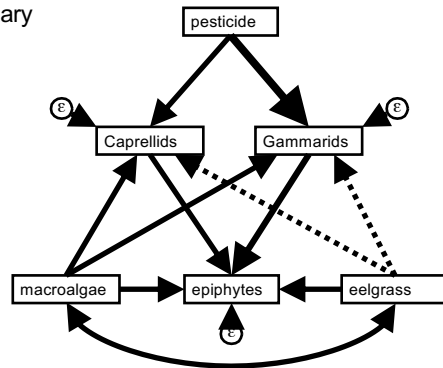
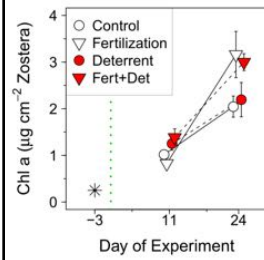
DEATH BY F-TABLES!

Notes: ANOVA tables for linear models describe the effects of chemical deterrent, nutrient fertilization, and sampling date on crustacean mesograzer density, gastropod mesograzer density, and epiphyte biomass. All data were natural-log-transformed except summer gastropods (square-root transformed). Model terms were tested using F tests and type III sums of squares. Note that the analyses presented for the summer experiment are balanced. P values <0.05 are shown in boldface.



From ANOVA to SEM. Increasing inference

Our model results imply that behind this summary of mean responses...



...is a network of effects like this.

From ANOVA to SEM. Increasing inference

- Teases out *complex relationships*
- Identification and comparison of *direct vs. indirect effects* & potential mediators
- Precise *mechanistic* explanations

