Structural equation modeling (SEM) in biology education research: moving beyond “what works” to understand how it works, and for whom.

Cissy Ballen¹*, Shima Salehi², & Sehoya Cotner³

¹Department of Biological Sciences, Auburn University, Auburn, AL
* mjb0100@auburn.edu
²Graduate School of Education, Stanford University, Stanford, CA
³Department of Biology Teaching and Learning, University of Minnesota, Minneapolis, MN

Geological Society of America Annual Meeting 2018
Indianapolis, Indiana, USA
The anxiety-performance link

Individual Environment (Social/Contextual)

Anxiety

Performance

Multiple individual and environmental factors account for the reciprocal relation between math anxiety and math performance.

Current Opinion in Behavioral Sciences

Chang and Beilock 2016
student deficit model
focuses on inadequacies of individuals
• assumes some students enter college lacking the academic resources necessary to succeed in an otherwise fair learning environment.
<table>
<thead>
<tr>
<th>student deficit model</th>
<th>course deficit model</th>
</tr>
</thead>
</table>
| focuses on inadequacies of individuals  
* assumes some students enter college lacking the academic resources necessary to succeed in an otherwise fair learning environment. | considers the negative impact of environmental conditions on student performance  
* assumes classroom practices favor certain groups of students while increasing performance disparities. |
student deficit model

change student behaviors

course deficit model

change classroom microclimate
student deficit model

change student behaviors

course deficit model

change classroom microclimate
Significant gendered performance differences in large introductory STEM lecture courses

Koester et al., 2016

These data show a correlation between GPA in other classes at Michigan and grade in STEM class. The graph indicates significant gendered performance differences, with women generally achieving lower grades compared to men. The dataset includes 627,998 students.
The microclimate of the classroom

Example: High stakes assessments

University of Minnesota
College Bio Sciences,
College Science &
Engineering (N = 1,562 )

<table>
<thead>
<tr>
<th>Class section</th>
<th>Instruct or</th>
<th>Class N</th>
<th>Women (%)</th>
<th>URM (%)</th>
<th>Average Age (SD)</th>
<th>Average ACT per class (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1</td>
<td>A</td>
<td>115</td>
<td>61.7</td>
<td>12.2</td>
<td>20.83 (2.47)</td>
<td>25.75 (3.42)</td>
</tr>
<tr>
<td>A.2</td>
<td>A</td>
<td>115</td>
<td>61.7</td>
<td>8.7</td>
<td>20.70 (2.03)</td>
<td>26.46 (3.55)</td>
</tr>
<tr>
<td>A.3</td>
<td>B</td>
<td>182</td>
<td>59.9</td>
<td>9.3</td>
<td>20.38 (2.52)</td>
<td>26.62 (2.81)</td>
</tr>
<tr>
<td>B.1</td>
<td>C</td>
<td>95</td>
<td>44.2</td>
<td>14.7</td>
<td>20.18 (3.15)</td>
<td>26.98 (3.81)</td>
</tr>
<tr>
<td>B.2</td>
<td>C</td>
<td>90</td>
<td>47.8</td>
<td>16.7</td>
<td>19.68 (1.70)</td>
<td>28.05 (3.17)</td>
</tr>
<tr>
<td>C.1</td>
<td>D</td>
<td>229</td>
<td>51.5</td>
<td>9.1</td>
<td>20.04 (2.16)</td>
<td>26.89 (3.80)</td>
</tr>
<tr>
<td>D.1</td>
<td>E,F</td>
<td>153</td>
<td>69.9</td>
<td>12.4</td>
<td>20.29 (2.41)</td>
<td>26.91 (3.55)</td>
</tr>
<tr>
<td>D.3</td>
<td>E,F</td>
<td>178</td>
<td>58.4</td>
<td>12.4</td>
<td>20.06 (1.86)</td>
<td>26.32 (3.53)</td>
</tr>
<tr>
<td>D.5</td>
<td>G</td>
<td>239</td>
<td>58.2</td>
<td>14.2</td>
<td>20.18 (2.19)</td>
<td>26.08 (3.55)</td>
</tr>
<tr>
<td>E.1</td>
<td>H</td>
<td>164</td>
<td>38.8</td>
<td>8.5</td>
<td>20.07 (1.96)</td>
<td>28.69 (3.50)</td>
</tr>
</tbody>
</table>
The microclimate of the classroom

Example: High stakes assessments

University of Minnesota
College Bio Sciences 2016
(N = 1,562 )

Question(s): 1) What is the extent of the gender gap in incoming academic preparation among students? 2) What is the extent of the gender gap in exam grades? 3) Do women and men report different levels of test anxiety and interest in science? 4) Do these two affective factors influence performance outcomes in undergraduate biology courses?
Question(s):
1) What is the extent of the gender gap in incoming academic preparation among students? 
   women’s ACT score on average is 0.28 standard deviation lower than men
Question(s):
1) What is the extent of the gender gap in incoming academic preparation among students?
2) What is the extent of the gender gap in exam grades?

Women’s exam scores on average was 0.15 standard deviation lower than men.
Question(s):
1) What is the extent of the gender gap in incoming academic preparation among students?
2) What is the extent of the gender gap in exam grades?
3) Do women and men report different levels of test anxiety and interest in science?

After controlling for students’ ACT, women reported 0.37 standard deviation higher interest in the science content, along with 0.43 standard deviation higher test anxiety than men.
Question(s):
1) What is the extent of the gender gap in incoming academic preparation among students?
2) What is the extent of the gender gap in exam grades?
3) Do women and men report different levels of test anxiety and interest in science?
4) Do these two affective factors influence performance outcomes in undergraduate biology courses?
• **Mediation** is a hypothesized chain in which one variable affects a second variable that, in turn, affects a third variable.

• The intervening variable, or the **mediator**, “mediates” the relationship.
ACT → Performance
Partial mediation model includes the direct effect of the ACT on performance.
full mediation model, whereby ACT score predicts performance indirectly through affective measures
Model fit indices

**Full mediation model for exam grade:**
\[
\chi^2 (8, N = 284) = 138.7, P < 0.0001
\]
root mean square error (RMSEA) = 0.385 (acceptable range: 0-0.07)
comparative fit index (CFI) = 0.144 (acceptable range: above 0.95)
standardized root mean square residual (SRMR) = 0.183 (acceptable range: 0-0.1).

**Partial mediation model for exam grade:**
\[
\chi^2 (2, N = 284) = 1.681, P = 0.431
\]
root mean square error (RMSEA) = 0.00 (acceptable range: 0-0.07)
comparative fit index (CFI) = 1 (acceptable range: above 0.95)
standardized root mean square residual (SRMR) = 0.027 (acceptable range: 0-0.1).
Model fit indices

**Full mediation model for exam grade:**
\[ \chi^2 (8, N = 284) = 138.7, P < 0.0001 \]
- root mean square error (RMSEA) = 0.385 (acceptable range: 0-0.07)
- comparative fit index (CFI) = 0.144 (acceptable range: above 0.95)
- standardized root mean square residual (SRMR) = 0.183 (acceptable range: 0-0.1).

**Partial mediation model for exam grade:**
- `GROUP analysis in Lavaan`
\[ \chi^2 (2, N = 284) = 1.681, P = 0.431 \]
- root mean square error (RMSEA) = 0.00 (acceptable range: 0-0.07)
- comparative fit index (CFI) = 1 (acceptable range: above 0.95)
- standardized root mean square residual (SRMR) = 0.027 (acceptable range: 0-0.1).
In our sample of women, incoming preparation (ACT) did not predict interest or anxiety, but both impacted exam performance.

A. Female students

**ACT**

- $b = -0.042$ (0.078)

**Interest**

- $b = -0.042$ (0.078)
- $b = 0.159^*$ (0.068)

**Exam Performance**

- Total effect = Direct Effect: $b = 0.551^{***}$ (0.064)

- $b = -0.113$ (0.079)
- $b = -0.218^{**}$ (0.068)

**Test Anxiety**

- $b = -0.127^*$ (0.059)
- $b = -0.297^*$ (0.116)

**Non-exam Performance**

- $b = -0.029$ (0.056)

- $b = 0.022$ (0.096)

- $b = -0.297^*$ (0.116)
- $b = 0.059$ (0.101)
In our sample of men, incoming preparation (ACT) predicted anxiety, but neither anxiety nor interest impacted exam performance.
Based on these data, we may expect the following to reduce gender gaps:

- Design curricula to promote students’ interests
- Minimize risk/threat during exams
- But - do these trends persist in upper division courses? In other disciplines?

Ballen et al. (2017)
Setting the scene...
Test Anxiety
Interest
ACT Exam Performance

Female students

\( b = -0.04 \)
\( P = 0.467 \)

Total effect = Direct Effect : \( b = 0.48 \)
\( P < 0.001 \)

\( b = 0.06 \)
\( P = 0.077 \)

\( b = -0.04 \)
\( P = 0.003 \)

\( b = -0.11 \)
\( P = 0.001 \)

\( b = 0.26 \)
\( P < 0.001 \)

CBS - lower

\( (N = 5,934) \)

Male students

\( b = 0.11 \)
\( P = 0.067 \)

Total effect = Direct Effect : \( b = 0.159^{***} \)
\( P < 0.001 \)

\( b = 0.04 \)
\( P = 0.443 \)

\( b = -0.061 \)
\( P = 0.271 \)

\( b = -0.43 \)
\( P < 0.001 \)
Test Anxiety
Interest
ACT Exam Performance

Total effect = Direct Effect: \( b = 0.48 \)
\( b = 0.06 \)

Female students

Test Anxiety
Interest
ACT
Exam Performance

\( b = -0.04 \)
\( P = 0.467 \)
\( b = 0.06 \)
\( P = 0.077 \)
\( b = -0.26 \)
\( P < 0.001 \)
\( b = -0.11 \)
\( P = 0.003 \)
\( b = -0.210 \)
\( P = 0.080 \)

Male students

Test Anxiety
Interest
ACT
Exam Performance

\( b = 0.11 \)
\( P = 0.067 \)
\( b = -0.43 \)
\( P < 0.001 \)
\( b = -0.061 \)
\( P = 0.271 \)
\( b = 0.08 \)
\( P = 0.909 \)
\( b = 0.01 \)
\( P = 0.954 \)

CBS - lower

CBS - upper
Test Anxiety
Interest
ACT Exam Performance

Total effect = Direct Effect: $b = 0.48$

$P < 0.0001$

Female students

ACT

Interest

Test Anxiety

Exam Performance

$P = 0.583$

$P < 0.0001$

$b = -0.04$

$b = -0.14$

$b = -0.26$

$P = 0.013$

$P = 0.0001$

$b = 0.03$

$P = 0.552$

$P = 0.044$

$b = 0.03$

$b = 0.11$

$b = 0.45$

$b = 0.48$

$P < 0.0001$

$b = 0.03$

$P = 0.552$

$P < 0.0001$

$b = 0.4$

$b = 0.04$

$b = -0.4$

$b = -0.04$

$P = 0.504$

$P = 0.554$

$b = -0.26$

$b = -0.14$

$b = -0.04$

$b = -0.04$

Test Anxiety

Exam Performance

CSE - lower
• What sort of effective interventions might mitigate these impacts?
<table>
<thead>
<tr>
<th></th>
<th>Low stakes</th>
<th>High stakes</th>
<th>Low stakes</th>
<th>High stakes</th>
<th>Low stakes</th>
<th>High stakes</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Biology 1003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Biology 3409</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**z-score**

- Men
- Women
student deficit model
focuses on inadequacies of individuals

course deficit model
considers the negative impact of environmental conditions on student performance
Acknowledgements – thank you for your time!

Special thanks to the National Association of Geoscience Teachers (NAGT) - GER Division

Statistical advice
Francoise Vermeylen
JD Walker

Test anxiety
Jonathan Andicoechea
Emma Goldberg
Will Harcombe
Mark Decker
Deena Wassenberg

Institutional Research and Planning
Marin Clarkberg

Undergraduate Researchers
Steven Wallace
Christine Lian
Olivia Treudeu
Mai Vang
Morgan Burkhart
Dahsol Lee