Supplemental Material

CBE—Life Sciences Education

Eddy *et al.*

Supplemental Materials

I. Example Course Materials: Guided Reading Questions

Students were encouraged to complete these before class-- as they read each chapter. They had a corresponding online homework assignment for the chapter to complete before class too.

Population Ecology

A. Population demographics

Define population density –

Define population distribution and describe 3 types of distribution patterns-

B. Survivorship- what does this mean?

Survivorship curves- what information can you learn by looking at one?

Draw each type (be sure to label your x and y axis). Describe each in words too and give example organisms that follow these patterns

Type I –

Type III –

Type II –

C. Population size and growth

Why would a population gain or lose individuals?

How is the per capita rate of increase for a population represented (symbol)? ____

How is it calculated?

Work through the example population of 100 rabbits: 100 rabbits in a field. 50 are born in one month to these 100 rabbits.

The net increase of rabbits per month is ___/100 = _____ = r

Exponential growth:

What is the equation for exponential population growth (G) =

So, if a rabbit population, in which r = 0.3 currently has 96 rabbits, what will the growth be in one month? (*Check your answer in Table 36.4)

Exponential growth model: how would you describe it in words?

How would you describe the exponential model as a curve plot? (Draw it).

Logistic growth model

What is a limiting factor? What does it do to growth?

What does a typical curve for logistic growth look like? Describe in words <u>and</u> draw it.

Define carrying capacity –

What is the equation for logistic growth?

As resources are depleted, what happens to the equation? That is, what happens to population growth as the population size approaches carrying capacity?

What is the difference between a density *dependent* and a density *independent* limiting factor?

D. Life History

If a population ecologist asks for your life history, what kind of information should you tell her?

Compare and contrast k-selection and r-selection

Where ideally on a logistic growth curve do we want to keep population size for say fish we hope to harvest for food and do not want to endanger?

E. Human population

What kind of growth (exponential or logistic) is demonstrated by the human population?

What year will the human population possibly "peak"?

What is the demographic transition?

What is the relationship between developed *vs.* non-developed countries to this transition? Describe in words what Figure 36.9B is illustrating:

F. Age structures:

Draw an age structure diagram for a hypothetical population with a high proportion of children and high birth rate. Be sure you label each axis and color code the information appropriately.

Even if methods are put in place with a population like you just drew above to reduce the birth rate (i.e. birth control), explain why stopping growth is like stopping a freight train.

Go to http://www.myfootprint.org/ and measure your ecological footprint. This is a measure of how your life affects the earth's limited resources. We'll talk about and compare these with each other in class.

So, how many planet earths would be needed if everyone lived like you? _______

(You may want to see the difference in your footprint as a college student at the University of North Carolina *vs*. your hometown life with your parents).

II. Controlling for Year to Year Variation in Exams and Student Ability

To attribute differences in student achievement to the change in course structure from low to moderate, we needed to establish exam equivalency and control for student variation between the low and moderate structure terms or, at least, control for those differences.

Controlling for Exam Differences. Exam equivalence has been established multiple ways in the literature ranging from the most rigorous (isomorphic questions; Rybarczyk, Baines, McVey, Thompson, & Wilkins, 2007) to establishing equivalency in question difficulty (Freeman, Haak, & Wenderoth, 2011) or equivalency in the cognitive skills requires to answer the questions (Crowe, Dirks, & Wenderoth, 2008). We employed the third approach. We used Bloom's Taxonomy to determine the cognitive skills that each exam required (*c.f.* Crowe et al. 2008). To minimize bias, we randomly combined exam questions from all 6 years into one master list. Two raters, blind to which term each question came from, independently determined the Bloom level of each question. After independently rating, they discussed any differences in scores assigned and reached a consensus score for all questions. A weighted Bloom Index was then calculated for each exam (*c.f.* Freeman et al. 2011). As each term had 4 exams, there were 12 exams under low structure and 12 exams under moderate. The Weighted Bloom Indexes of the exams were compared using a Kruskal-Wallis test to determine if the academic challenge of the exams had changed across treatments.

Results. The weighted Bloom Index for each exam did not differ significantly under low and moderate structure (Kruskal-Wallis test: H=.274, d.f.=2, p=0.6). The median weighted Bloom Index was a 0.352 (±0.03 sd). On average, 32 ± 6.0% of the exam questions where at higher Bloom levels (Application High, Analysis, Synthesis and Evaluate). Overall, we concluded that the exams were homogeneous across treatments.

Controlling for Previous Student Academic Achievement. This study took place over four years. Term to term variation in student academic ability is not uncommon, so be confident that any gains we observed were due to the intervention and not changes in the student body, we controlled for student academic ability. To control for differences in student academic ability, we collected student data on metrics known to be correlated with student performance in college classes from a term not included in the study (N = 551). Unfortunately, the data available to us was limited as our students were primarily first term first-years. We collected variables from the Registrar that would be available upon entrance to college including: High school GPA, SAT I Reading, SAT I Math, SAT I written, and SAT 2 exam scores, ACT scores, gender, whether or not the student was pursuing a science degree, racial and ethnic group and student age. High school GPA was a problematic predictor variable as AP credit was scored differently from school to school. With AP credit it was possible for students to score above a 4.0 and how much credit above a 4.0 a student could earn for each AP class varied from school to school. Thus, instead of using the actual GPA we used a high school GPA over 4.0 as a binary indicator that a student was able to take AP courses in high school and earned at least a B in the course. In addition, few students took either SAT 2 or ACT exams, so we could not use these metrics without having to throw out the majority of our sample size. For students who only took the ACT, we were able to use a conversion scale provided by ACT, Inc. to convert their scores to the SAT I scale. This SAT I scale was based on a combined reading and math score, so the SAT I Math and Reading exams were combined for all students. Only students who we had data on all these measures for were included in the final data set.

We used a powerful multi-model inference techniques implemented in R using the package MuMIn (Barton, 2013) to determine which combination of the remaining variables (SAT I combined score, HS.GPA greater than 4.0, age, whether or not pursing a science degree, and gender) best predicted student grade for a term of the course not included in the study itself. Model selection is a superior method for exploratory analyses because it can(1) identify the best model for predicting a response variable (as well as quantifying how confident one can be that the best model is actually the best), (2) calculate regression coefficients that incorporate information from multiple models (incorporating model uncertainty), and (3) determine the importance of individual variables relative to one another for predicting the response variable (Burnham & Anderson, 2002). Information theoretic model selection does not rely on p-values like traditional forward or backward model selection, but instead focuses on a measure of the information to noise ratio for each model, the Akaike Information Criterion corrected for small sample sizes (AICc; Akaike, 1974; Garamszegi, 2011; Burnham, Anderson, & Huyvaert, 2011). The smaller the AICc value the more information is present in the model (Burnham & Anderson, 2004). This measure is calculated for each model and the best model (the one with the smallest AICc value) is used as a reference to compare the rest of the models to. A change in AICc (Δ_i) > 10 implies that the model contains practically no information (Burnham & Anderson, 2004). Finally, we also calculate model weights; a measure of the likelihood of the observed data given the model used that have been standardized to add up to one. These weights can be used to compare models as the likelihood is approximately the probability that the model is actually the best model. We also used the Akaike weights to calculate a measure of the relative importance of individual variables. This process involved summing the Akaike weights across all the models that include a particular explanatory variable. The relative variable importance gives us a measure of how important the variable was likely to be (probability that a certain variable is important for explaining variation in student performance). The variables identified via model selection were then included in all subsequent models to account for variation in student ability within and between years.

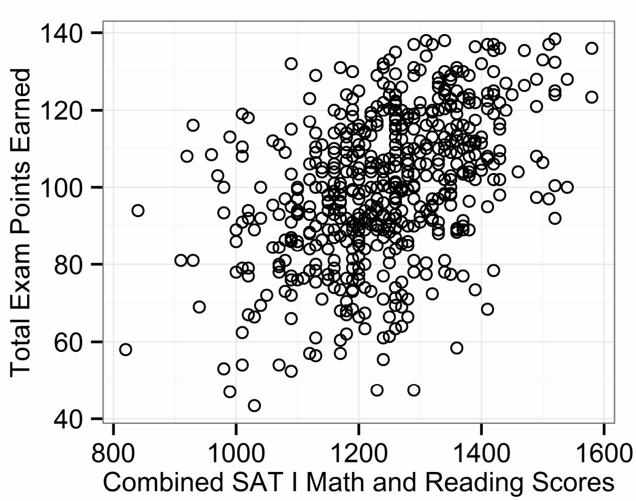
Results. Our initial model to determine which variables best predicted student performance involved 569 students. The initial full model included a student's combined SAT I Math and Reading scores, student age, whether or not student was majoring in science, student gender and whether or not their high school GPA was over 4.0. Using model selection, we found 16 reasonable (Δ_i<10) regression models to predict student performance. The best model (R^2=.193) included only the combined SAT I Math and Reading exam scores and intercept as predictors (Table S1; Figure S1). The combined SAT I Math and Reading scores was also the only variable with a relative importance of 1 (i.e. was present in every model in the subset of best models). The variable with the next highest relative importance was whether or not a student's high school GPA was over 4.0 (present in 42% of the best models). The first model including only combined SAT I scores was 1.5 times as likely as the second best model that included combined SAT score and whether or not a student's high school GPA was over 4.0. Averaging across all models, the combined SAT I score was the only predictor with a β significantly different from zero (β=0.069±0.0061 se, $p < 0.0001$, Table S2). Given these findings, of the predictors for student grade we started with we chose to include only the combined SAT I Math and Reading scores in subsequent models.

Supp. Table 1. Variables influencing student performance in an introductory biology course for mixed majors at UNC Chapel Hill. Only models that are most informative (Δi<4) are shown. Models are ranked from those with the most support to those with the least.

	Model	R^2	Df	logLikelihood	AICc	Delta	Weight
1	Int + SATI.Comb	0.193	3	-2410.08	4826.2	0	0.27
2	Int + SATI.Comb + Over4.0	0.195	4	-2409.49	4827.1	0.84	0.18
3	Int + SATI.Comb + Gender	0.194	4	-2409.85	4827.8	1.57	0.12
4	Int + SATI.Comb + Sci.Degree	0.194	4	-2410.06	4828.2	1.98	0.10
5	Int + SATI.Comb + Over4.0 + Gender	0.196	5	-2409.14	4828.4	2.17	0.09
6	Int +SATI.Comb + Over4.0 + Sci.Degree	0.196	5	-2409.43	4829.0	2.76	0.07
7	Int + SATI.Comb + Sex + Sci.Degree	0.194	5	-2409.84	4829.8	3.58	0.05

Supp. Table 2. Relative parameter importance for student performance in an introductory biology course for mixed majors at UNC Chapel Hill across all models with a $\Delta i<10$ as well as averaged model parameters.

Parameter	Relative variable importance	Model averaged β ± adjusted SE	p-value
Intercept:	NA	13.27±7.60	0.082
SATI.Comb:	1	0.0688±0.00606	$<2\times10^{-16}$
Over.4.0: Yes	0.42	2.801±2.422	0.248
Gender: Male	0.32	1.19±1.61	0.462
Sci.Degree: Yes	0.27	0.364±1.51	0.810
Age (ref =18): 17	0.09	-5.072±10.36	0.625
19		-0.031±1.740	0.986
20		-2.41±2.365	0.308
21		3.982±4.017	0.323
22		3.497±5.268	0.508
23		26.67±12.60	0.035
27		23.93±17.83	0.181



Supp. Figure 1. Combined Math and Verbal SAT scores and Total Exam Points.

III. Correlations between survey question responses

During the low structure term, student response level for hour spent studying was correlated with 8 of the 9 remaining explanatory variables and reported frequency with which a student reviewed their notes after class was correlated with all 9 other variables. The four strongest correlations were a positive correlation between reviewing notes after class and spending more hours studying (0.476, $p < 0.0001$), reviewing notes after class and speaking during class (0.458, $p < 0.0001$), the number of hours a student spent studying and seeking extra help (0.402, $p < 0.0001$), and speaking out in class and seeking extra help. During the two intervention terms, the highest correlations were a positive correlations between hours spent studying and frequency student reviewed notes (0.483, $p < 0.0001$), speaking in class and seeking extra help (0.435, $p<0.0001$), and frequency of completing the practicing questions and visiting the textbook's website (0.348, $p < 0.0001$).

Supplemental Table 3. Gamma Rank Correlations between level of responses to Survey Questions about study habits during the low structure (below the diagonal) and moderate structure terms (above the diagonal). Bolded correlations were significant after correcting for the false discovery rate. Estimated p-values are in parentheses (based on 1,000 iterations of a randomization test).

	Hours Study	Speaking in Class	Sense of Comm.	Review Notes	Read before Class	Complete Practice Questions	Visit Textbook Website	Form Study Group	Seek Extra Help	Attend Lecture
Hours Study		**0.243** (2.2e-16)	**0.131** (0.014)	**0.483** (2.2e-16)	**0.293** (2.2e-16)	**0.128** (0.008)	-0.0864 (0.117)	**0.123** (0.009)	**0.334** (2.2e-16)	NA
Speaking in Class	**0.27** (0.004)		**0.265** (2.2e-16)	**0.338** (2.2e-16)	**0.237** (2.2e-16)	**0.173** (2.2e-16)	**0.141** (0.003)	**0.168** (2.2e-16)	**0.435** (2.2e-16)	NA
Sense of Comm.	.159 (0.05)	0.131 (0.09)		**0.168** (0.001)	**0.115** (0.016)	**0.140** (0.003)	0.0119 (0.815)	**0.300** (2.2e-16)	**0.211** (2.2e-16)	NA
Review Notes	**0.476** (2.2e-16)	**0.458** (2.2e-16)	0.110 (0.103)		**0.175** (2.2e-16)	**0.338** (2.2e-16)	**0.213** (2.23-16)	**0.155** (0.003)	**0.293** (2.2e-16)	NA
Read	**0.289** (2.2e-6)	**0.296** (2.2e-16)	0.077 (0.103)	**0.354** (2.2e-16)		0.029 (0.524)	0.0503 (0.233)	0.0321 (0.465)	**0.168** (2.2e-16)	NA
Practice Questions	**0.258** (2.2e-16)	**0.187** (0.014)	**0.178** (0.006)	**0.283** (2.2e-16)	0.0675 (0.275)		**0.348** (2.2e-16)	**0.170** (0.001)	**0.215** (2.2e-16)	NA
Visit Website	**0.258** (0.005)	0.115 (0.186)	0.0973 (0.214)	**0.359** (2.2e-16)	-0.158 (0.047)	**0.372** (2.2e-16)		**0.223** (2.2e-16)	**0.222** (2.2e-16)	NA
Study Group	**0.198** (0.0009)	**.326** (2.2e-16)	0.0986 (0.131)	**0.282** (2.2e-16)	-0.0118 (0.83)	0.0858 (0.174)	0.032 (0.647)		**0.209** (2.2e-16)	NA
Extra Help	**0.402** (2.2e-16)	**.451** (2.2e-16)	0.0810 (0.269)	**0.359** (2.2e-16)	**0.158** (0.021)	**0.207** (0.003)	**0.306** (0.001)	**0.303** (2.2e-16)		NA
Attend Lecture	0.0290 (0.617)	0.0187 (0.741)	-0.0585 (0.231)	0.00976 (0.841)	0.00709 (0.995)	-0.0570 (0.244)	0.0665 (0.257)	**-0.123** (0.004)	0.210 (0.704)	

IV. Survey Questions

1. In a typical week, how many hours did you spend studying or preparing for this course?

a) 0
b) 1-3
c) 4-7
d)7-10
e) > ten hours

2. How often did you come to class without completing the Reading?

a) Often
b) Sometimes
c) Rarely
d) Never

3. How important were the *homework assignments* to your overall understanding of course material?

a) Very important
b) Important
c) Somewhat important
d) Not at all important
e) Did not complete

4. How often did you review your lecture notes after each class?

a) Often
b) Sometimes
c) Rarely
d) Never

5. How often did you answer the practice questions at the end of each chapter?

a) Often
b) Sometimes
c) Rarely
d) Never

6. How often did you ask a question or make a comment that contributed to class discussions?

a) Often
b) Sometimes
c) Rarely
d) Never

7. How often did you work with classmates outside of class on assigned work or studying?

a) Often
b) Sometimes
c) Rarely
d) Never

8. Students in this class know each other.
a) Strongly agree
b) Agree
c) Neutral
d) Disagree
e) Strongly disagree

9. Students in this class help one another (e.g. sharing notes when absent, etc.).
a) Strongly agree
b) Agree
c) Neutral
d) Disagree
e) Strongly disagree

10. Students in this class think of themselves as a community.
a) Strongly agree
b) Agree
c) Neutral
d) Disagree
e) Strongly disagree

11. How much of your coursework emphasized *memorizing* facts, ideas or methods from lecture and readings so that you can repeat them in the same form.
a) Most
b) Quite a bit
c) Some
d) Very little
e) None

12. How often did you attend lecture?
a) Often
b) Sometimes
c) Rarely
d) Never

13. I will likely use study skills learned in this class in future courses.
a) Strongly agree
b) Agree
c) Neutral
d) Disagree
e) Strongly disagree

14. How important was the *lecture component* of this course to your overall understanding of course material?
a) Very important
b) Important
c) Somewhat important
d) Not at all important
e) Did not attend

15. How old are you?
a) 0-18 years old
b) 19-20 years old
c) 21-22 years old
d) 23 years and older

16. I am a:
a) Male
b) Female

17. I am a:
a) First year
b) Sophomore
c) Junior
d) Senior
e) Continuing Studies/Other

18. Which of the following best describes your parents' education:
a) Both parents attended college
b) Only father attended college
c) Only mother attended college
d) Neither parents attended college
e) I don't know

Supplemental Materials References

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, *19*(6), 716-723.

Bartoń, K. (2012). MuMIn: multi-model inference. *R package version*, *1*(2).

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: a practical information-theoretic approach*. Springer.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociological methods & research*, *33*(2), 261-304.

Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, *65*(1), 23-35.

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in bloom: implementing Bloom's taxonomy to enhance student learning in biology. *CBE-Life Sciences Education*, *7*(4), 368-381.

Freeman, S., Haak, D., & Wenderoth, M. P. (2011). Increased course structure improves performance in introductory biology. *CBE-Life Sciences Education*, *10*(2), 175-186.

Garamszegi, L. Z. (2011). Information-theoretic approaches to statistical analysis in behavioural ecology: an introduction. *Behavioral Ecology and Sociobiology*, *65*(1), 1-11.

Rybarczyk, B. J., Baines, A. T., McVey, M., Thompson, J. T., & Wilkins, H. (2007). A case-based approach increases student learning outcomes and comprehension of cellular respiration concepts. *Biochemistry and Molecular Biology Education*, *35*(3), 181-186.