Supplemental Material

CBE—Life Sciences Education

Wright *et al.*

Table A: The absolute number of low-SES and middle/high-SES students broken down by race/ethnicity/nationality and gender.		
Student Race/Ethnicity	**Middle/High-SES (n = 3993)**	**Low-SES (n = 817)**
White/Caucasian	2039	55
Asian	1421	344
Black / African American	12	108
Hawaiian / Pacific Islander	7	36
International	327	0
Hispanic / Latin@	31	224
Native American	12	44
NA	144	6
Gender	**Middle/High-SES (n = 3993)**	**Low-SES (n = 817)**
Male	1728	292
Female	2265	525

Table B: Rubric used for determining the difficulty of each question on an exam.	
Instructions for Analysts: Rate question difficulty on a scale of 1-3. Both analysts should come to consensus on each question, and BLOOM criteria should not be used. Use percentages to code if stuck between two answers and to come to consensus. Record the number of responses available to the student.	
Easy Questions (1) 80% or more students will answer correctly, awarded 75% or more of points if partial credit is awarded	**Hard Questions (3)** 60% or fewer of the students will answer correctly, or get 50% or fewer points if partial credit is awarded
How the question is written: Question stem or answer choices **uses similar or identical language to class lectures or materials**. Question **strongly hints at answer** (lays out the logic of the question or gives details that makes questions easier)	**How the question is written:** **Reading Comprehension is challenging** (unclear, poorly worded or otherwise difficult question stems) **Detailed questions stems with distractors** (students need to sort out irrelevant information, or information is included that could be misapplied) Questions that present the concept from an **atypical point of view** (especially if this differs from how the concept is presented in lecture or class materials) Multiple choicequestions that have an **"all of the above"** choice or questions that ask "**what is more true?"** (score a 2 or above)
Details: Questions that ask a **detailed** question about an **example** used to explain a larger concept covered in class.	**Details:** Questions asking students to **extrapolate larger concepts from detailed information**. Can score a 2 if these larger concepts have been emphasized in class.
Number of concepts: **Single concepts** discussed with a **significant amount of class time spent** (may have associated activity, at least one example	**Number of concepts:** Questions bringing together **multiple concepts from class**, esp. in ways that the students have discussed little (score a 2 if

discussed in addition to general concept)	this is discussed)
Recall: Questions requiring **recall of emphasized details** (especially concepts where few details are taught)	**Recall:** Very **specific recall about a concept detail**, mentioned once or only included on class notes w/o verbal mention, especially if the detail is from a topic where a lot of detailed information is also taught.
Topic: Question tests a **straightforward** topic that may be encountered without a previous background of science classes or that follow "common sense" logic.	**Topic:** Question **tests known difficulties or misconceptions,** or topics that are intrinsically difficult for students to grasp. Can score a 2 if addressed in class.
Calculations: **Simple calculations** with provided equations **Multiple equation** questions, where the process of solving these types of questions has been **well-emphasized** in class/lab	**Calculations:** Questions with **multiple equations**, with multiple equations needing to be used, and/or no specification of which need to be applied Can rate higher if what **variables** are used and where students should use them are more **ambiguous**
Lab: Concepts **covered in lab** (big take-home messages, can be scored as a 2 for more challenging labs)	**Lab:** Large, thematic concepts that are **not reinforced by lab** (score a 2 if considerable time is spent in class)
System thinking: **"System" questions** that address topics that are **well-emphasized,** ask for little more than **recall**, require few details or synthesis of multiple concepts, or	**System thinking:** **"System"** thinking that has **students respond to perturbations** in a biological system or systems in different contexts. More difficult questions ask for greater detail or effects several steps downstream.
Experimental Design: **Experimental design** questions are easy if they are more geared towards **recall** of classic experiments, especially those gone over in detail.	**Experimental Design:** **Experimental design** questions, these ask students to **set up** experiments ask students to demonstrate knowledge of how to design a specific, controlled experiment.

Graphing: **Graphing** questions that **provide labels**, and what the **format** of the graph should be (bar graph, scatter plot, etc). Graphs may be similar to those presented in class.	**Graphing:** **Graphing** questions that do not provide labels, ask the student must determine the best format to use, and ask the student to include a statistical component (error bars, linear regression), or present a novel situation for analysis.
Drawing and Pictures: **Labeling** simple diagrams, or **drawing** with **few components** to include. Does not ask for recall of many details.	**Drawing and Pictures:** **Labeling** complicated diagrams or those with many components, **detailed drawings** requiring recall of many pieces of information.

Table C: Correlation matrices output for numerical fixed-effects[a,b]				
	Cum.GPA	Percent CR	W.Blooms	W.Diff
Cum.GPA	1.00	-0.042	-0.040	0.025
Percent CR	0.042	1.00	**0.47**	-0.016
W.Blooms	-0.040	**0.47**	1.00	0.10
W.Diff	0.025	-0.016	0.10	1.00

[a]Cum.GPA = cumulative college GPA at start of introductory biology series; Percent CR = percentage of constructed-response questions on an exam; W.Blooms = weighted Bloom's level of an exam; W.Diff = weighted difficulty of an exam

[b]Bolded values have correlation values that are $x > 0.50$ or $0.50 > x > 0.30$.