# Increased Course Structure Enables Instructors to Increase Introductory Biology Exam Rigor

Adriana J. LaGier

Grand View University, College of Social and Natural Sciences, Department of Biology, 1200 Grandview Ave, ELS206, Des Moines, IA 50316; Tel: (515) 263-2874; Email: alagier@grandview.edu

**Abstract**

In courses with a heterogeneous student population, instructors are often challenged to balance successful course completion with rigor. This difficult task can be confounded in foundational, gateway courses, such as introductory biology, which serves a mix of freshman majors at various levels of preparedness. Research suggests that changes in course design, such as increasing course structure, can offer a solution. We hypothesize that increased course structure enables instructors to increase exam rigor without coincidentally increasing failure rates. Sixteen sections of general biology classes over the course of eight semesters were analyzed; eight sections had relatively low class structure (i.e., mostly Socratic learning and clickers), while eight sections had moderate structure (i.e., including class note summaries and practice exams). Weighted Bloom's Index of 150 exam questions was used to facilitate comparisons between designs. Although exam rigor increased (as gauged by WBI), in moderate structured courses, student exam scores, perception of the difficulty of the subject matter and failure rates did not change in comparison to low structured courses. This study supports the use of increased course structure to balance student success and rigor. Additionally, it supports the use of Weighted Bloom's Index as a method for assessing exam equivalence across institutions.

**Keywords:** general biology, course structure, exam rigor, weighted Bloom's taxonomy

## Introduction

For teachers, the increased access to higher education means that there is a high variability in student preparedness for college. In the 2000s, only 51% of high school student who took the ACT college entrance exams are ready for college-level reading (ACT, 2006). In courses with such a heterogeneous student population, instructors are often challenged to balance successful course completion with rigor. This difficult task can be confounded in foundational, gateway courses, such as introductory biology, which serves a mix of freshman majors at various levels of preparedness.

Research suggests that simple changes in course design, such as increasing course structure, can offer a solution (Freeman et al., 2011). Although a course's rigor can be defined at multiple levels (e.g., fast-paced, high degree of time and energy), this study focused on cognitive expectations, particularly the depth of question asked during summative assessment such as exams. To help instructors qualify the rigor of their exams, Freeman and colleagues developed the weighted Bloom's index scale (WBI). Bloom's taxonomy of learning identifies six levels of understanding any topic (Bloom, 1956). The WBI summarizes the average Bloom's level of exam questions weighted by the points possible.

Increases in the index indicate an increase in higher-order cognitive skills and are associated with increased cognitive complexity (Anderson and Krathwhol, 2001). Higher level cognitive processes are a criterion by which rigor is defined (Wyse and Soneral, 2018). Thereby, the WBI is a valuable tool in assigning a value of rigor to an exam and by association to a course.

Increasing rigor has been shown to increase student engagement (Paige et al., 2013). Students rise to the occasion as long as they perceive that they are supported in their endeavor (Adams, 2020). Teachers unnecessarily worry that increasing rigor will lead to increased failure rates (Attis, 2016). Indeed, there is a low correlation between a student's perception of rigor and their learning (Duncan et al., 2013).

The aim of the study presented here was to determine whether increasing course structure enables an instructor to increase exam rigor without coincidentally increasing failure rates.

| Course Structure Format | Low | Moderate |
|---|---|---|
| Lecturing | √ | √ |
| Thin-pair-Share Activities | √ | √ |
| Clickers | √ | √ |
| Practice Exams | | √ |
| Class Note Summaries | | √ |
| In-class Group Activity | | √ |

**Table 1:** Criteria Defining Low/Moderate Course Structure

**Materials and Methods**

Sixteen sections of an introductory (for majors) general biology (BIOL101) class were analyzed; eight sections had relatively low class structure, while eight sections had moderate structure (see table for criteria defining structure). Classes were taught over the course of eight semesters; two sections taught per semester by the same instructor.

Class size and majors data were collected using TurningPoint® (Turning Technology), i.e., clickers. Over the eight semesters in which the classes were taught, no significant patterns were noted in either parameter. Grand View University, where these classes were taught, is a primarily undergraduate university with a 97% acceptance rate with a total enrollment hovering at about 1,800 students (~85% of full-time students). Class size for both course structures was 23 ± 3 students (N=8 course sections for each).

Weighted Bloom's Index of each exam was used to facilitate comparisons between designs. 150 exam questions (fill-in and multiple choice) per semester separated over five exams given progressively through the semester were indexed. Each exam question was classified according to the complexity of the mental processes involved and assigned a rank based on Bloom's taxonomy of learning (knowledge, comprehension, application, analysis, synthesis and evaluation). Ranks (1-6) were used to calculate the index as follows (where P is points per question, B is Bloom's rank, T is total points):

$$Weighted\ Bloom's\ Index = ((\sum_{1}^{n} P * B) \,/\, (T * 6)) * 100$$

To further clarify, a student taking a low course structure class (Student Low) would expect to spend the majority of time in class listening to a lecture. Periodically, they would participate individually in a clicker question and as a group in a think-pair-share activity. Alternatively, a student taking a moderate course structure class (Student Mod) would expect to have the same experience as a student in a low course structure class. However, "student mod" would be provided with class note summaries and participate in a bi-weekly in-class group activity (e.g., codon bingo). Approximately every three weeks, "student mod" would participate in an in-class group practice exam (note: optional with no points).

In both low and moderate course structured sections, a multiple choice exam was scheduled approximately every three weeks for a total of five semester exams. "Student low" would take four exams with a WBI of <30 (i.e., indicating mostly recall) and one exam with a WBI of ~30 (i.e., indicating an increase in higher order cognitive skills). "Student mod" would take four exams with a WBI of >30 (i.e., indicating mostly conceptual questions) and one exam with a WBI of ~30 (i.e., indicating mostly recall).

Student perceptions of course were evaluated using IDEA diagnostic form reports (IDEA student ratings system, Kansas City University). This evaluation system poses 40 multiple choice questions to the students asking them to evaluate their progress on relevant course learning objectives, instructor teaching methods and overall impression of the instructor and course. Failure rate data was collected by instructor with Student Success Collaborative (SSC) software (educational advisory board, EAB) providing historical data.

Statistical analysis using one-way ANOVA supplemented with post-hoc Tukey HSD multiple comparison was performed with Astatsa statistical calculator.

**Results**

In a moderate course structure class increasing Weighted Bloom's Index did not alter exam grades in comparison to a low course structure class. In the general biology lecture course, five exams were taken progressively throughout the fifteen-week semester. Each

exam was assessed by Weighted Bloom's Index (WBI). WBI ranges from 0-100 and progresses from lower-order to high-order cognitive skills. The average ± SD WBI when considering all five exams was 24.8 ± 3.6 for low versus 33.9 ± 3.5 for moderate course structure. The WBI of low course structure indicates mostly recall, while the WBI of moderate course structure indicates mostly conceptual. WBI for each of the five exams administered to low and moderate course sections had a WBI ranging from 22 to 38 with 30 being the median (Figure 1 bottom).

The mean grade was calculated for each exam for low and moderate course structure (average ± SD for 8 class sections each). The average ± SD exam grade when considering all five exams was 71.9 ± 3.5 for low and 72.1 ± 1.3 for moderate course structures. Overall, no significant difference in exam performance was noted between low versus moderate structured courses for exam 1, 2 and 5 where exam WBI was tailored for the corresponding course structure. (Figure 1 top).

To evaluate whether course structure was linked to type of exam administered was increased for one exam (exam #3) in a low course structure class and WBI was decreased for one exam (exam #4) in a moderate course structure class This design disrupted the association of low course structure sections being administered low WBI exam and vice versa. The increase in an exam's WBI to mostly conceptual questions was associated with a significant dip in exam performance in low course structure in comparison to moderate course structure. Whereas the decrease in an exam's WBI to mostly recall did not significantly alter exam performance in moderate course structure in comparison to low course structure (Figure 1 top).

Student perception of the difficulty of the subject matter not altered by increased exam rigor. Sixteen sections of general biology classes over the course of eight semesters were analyzed; eight sections had relatively low structure (i.e., mostly Socratic learning and clickers), while eight sections had moderate structure (i.e., including class note summaries and practice exams). Student perception of difficulty of the subject matter assessed by IDEA diagnostic report using a 5-point scale (1 being much less and 5 being much more than most courses) showed no significant difference (Figure 2).

Change in course structure did not alter failure rates. Exams points made up >50% and no more than 65% of any class. Classes were composed of 23 ± 3 students. The percentage of

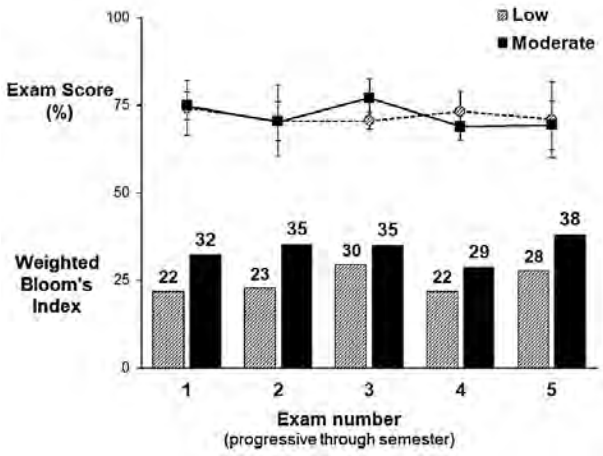


**Figure 1:** Relationship between Each exam's Weighted Bloom's Index in low and moderate course structure sections with mean exam score.

Five exams were taken progressively throughout the semester. **Bottom:** Each exam was assessed by Weighted Bloom's Index (WBI); low (diagonal stripe bar) and moderate (black bar) course structure. WBI ranges from 0-100 and progresses from lower-order to high-order cognitive skills. Class size for both course structures was 23 ± 3 students (N=8 course sections for each). **Top:** The mean grade was calculated for each exam for low (dotted line with circles) and moderate (solid line with squares) course structure (average ± SD for 8 class sections each).

Exam number 3 = $p < 0.05$ low course structure vs. moderate course student mean exam score.

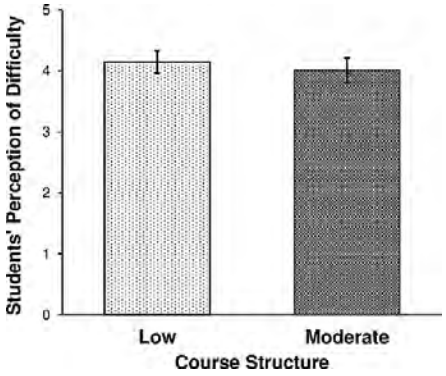


**Figure 2:** Although exam rigor increased in moderate structured classes, student perception of the difficulty of the subject were not altered.

Sixteen sections of general biology classes over the course of eight semesters were analyzed; eight sections had relatively low structure (i.e., mostly Socratic learning and clickers) [lightly dotted bar], while eight sections had moderate structure (i.e., including class note summaries and practice exams) [heavily dotted bar]. Student perception of difficulty of the subject matter assessed by IDEA diagnostic report using a 5-point scale (1 being much less and 5 being much more than most courses). Classes composed of 23 ± 3 students. Bars are average ± standard deviation; *N=8 sections.*

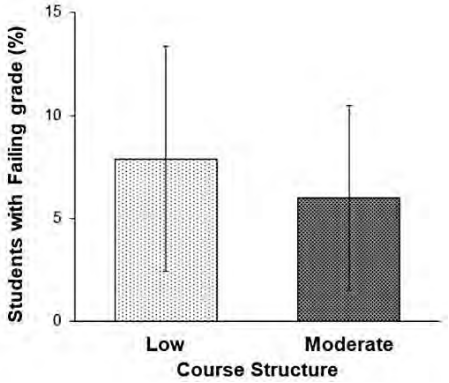


**Figure 3:** Although exam rigor increased in moderate structured classes, student failure rates were not altered.

Sixteen sections of general biology classes over the course of eight semesters were analyzed; eight sections had relatively low structure (i.e., mostly Socratic learning and clickers) [lightly dotted bar], while eight sections had moderate structure (i.e., including class note summaries and practice exams) [heavily dotted bar]. Percentage of students in each section who were assigned a grade of "F". Classes composed of 23 ± 3 students. Bars are average ± standard deviation; *N=8 sections.*

students in each section who were assigned a grade of "F" (defined by obtaining less than 60% of the overall points for the course) showed no significant difference between low and moderate structure classes (Figure 3).

**Conclusions**

The data presented here supports the study aim that increased course structure in introductory general biology courses helps balance student success and rigor as judged by cognitive expectations on summative assignments, i.e., exams. Without fearing impacting failure rates, a teacher can effectively increase the rigor of the course, particularly increasing conceptual skill questions on exams, by making small course structure adjustments such as adding class summaries, group activities and practice exams.

Students in low structured course sections challenged with a higher WBI exam showed a significant dip in exam performance in comparison to students in moderate structured course sections. This data supports the idea that exam rigor should only be increased with an associated increase in course structure. Alternatively, students in moderate course sections administered a lower WBI exam showed no significant change (positive or negative) in exam performance in comparison to students in low structured course sections. Although this result will require further investigation into student mindset., it does support the idea that WBI could be increased beyond the 35 in moderate structure courses.

Beyond summative assessments, designing a moderate course structure fortuitously increases other forms of rigor. Pacing, time and energy are alternate ways of defining rigor (Winston et al., 1994). In order to practically conduct in-class mock practice exams, class content, which stayed the same between low and moderate course structures, was presented at a faster-pace. In addition, because group activities actively engage students, they require more energy from the students and studying class summaries require more time. Yet, this study shows that increasing rigor in these myriad of ways does not necessarily impact student perceptions of the course rigor. Therefore, modifying an existing course to include moderate course structure should help teachers

increase course rigor without fearing reducing student success.

Additionally, this study supports the use of Weighted Bloom's Index (WBI) as a method for assessing exam equivalence across semesters and across institutions. A teacher could practically use WBI calculations of exams as a tool to gauge exam rigor. However, caution should be taken to consider other factors when ranking exam questions. Current research indicates that students rank Bloom's questions differently than expected. Often times, these perceptions are driven by questions that unintentionally increase cognitive load (Phillips et al., 2019). For example, a question can add unnecessary information and/or complex scenarios to appear rigorous and yet, the intent is recall or vice versa. In future studies utilizing WBI, cognitive load factors in the design of the exam question should be considered.

**Acknowledgments**

Grand View University for supporting the use of clickers, IDEA and providing access to the SSC software. Department of Biology for scheduling support and Grand View University students for providing course evaluations

**References**

ACT. (2006). Reading Between the Lines: What the ACT Reveals About College Readiness in Reading, Iowa City.

ADAMS, S., (2020). Teacher and Student Perceptions of Rigor and Relevance in Grades 6-8. School of Education of Baker University Thesis Defense.

ANDERSON, L. W. & KRATHWOHL, D. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman.

ATTIS, D. (2016). The false tradeoff between academic rigor and course pass rates. EAB insights. https://eab.com/insights/blogs/student-success/the-false-tradeoff-between-academic-rigor-and-course-pass-rates/.

BLOOM, B. S., ENGELHART, M. D., FURST, E .J., HILL, W.H., & KRATHWOHL, D. R. (1956). Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook. Cognitive Domain, New York: David McKay.

DUNCAN, H. E., BRET, R. & HVIDSTON, D. (2013). Exploring student perceptions of rigor online: toward a definition of rigorous learning. Journal on Excellence in College Teaching, 24(4):5.

FREEMAN, S., HAAK, D., & WENDEROTH, M.P. (2011). Increased course structure improves performance in introductory biology. CBE-Life Science Education, 10:175.

PAIGE, D. D., SIZEMORE, J. M., & NEACE, W. P. (2013). Working Inside the Box: Exploring the Relationship Between Student Engagement and Cognitive Rigor. NASSP Bulletin, 97(2):105. doi: 10.1177/0192636512473505.

PHILLIPS, A., BRIGGS, J. & JENSEN, J. (2019). Beyond Bloom's: Students' Perception of Bloom's Taxonomy and its Convolution with Cognitive Load. Journal of Psychological Research. 1. doi: 10.30564/jpr.v1i01.421.

WINSTON, R. B., VAHALA, M. E., NICHOLS, E. C., GILLIS, M. E., et. al. (1994). A measure of college classroom climate: The College Classroom Environment Scales. Journal of College Student Development, 35(1), 11.

WYSE, S. A. & SONERAL, P. A. G .(2018). "Is this class hard?" Defining and analyzing academic rigor from a learner's perspective. CBE-Life Science Education, 17(4):ar59,1-14. doi: 10.1187/cbe.17-12-0278.