UC Davis
UC Davis Previously Published Works

Title

Dynamic answer-dependent multiple-choice questions and holistic assessment analysis in high-enrollment courses.

Permalink

Journal

ISSN

Authors

Atwal, Harnejan

Quides, Kenjiro

Publication Date

DOI

Dynamic answer-dependent multiple-choice questions and holistic assessment analysis in high-enrollment courses

Harnejan K. Atwal,[1] Kenjiro W. Quides[1,2]

AUTHOR AFFILIATIONS See affiliation list on p. 4.

ABSTRACT   Many 4-year public institutions face significant pedagogical challenges due to the high ratio of students to teaching team members. To address the issue, we developed a workflow using the programming language R as a method to rapidly grade multiple-choice questions, adjust for errors, and grade answer-dependent style multiple-choice questions, thus shifting the teaching teams' time commitment back to student interaction. We provide an example of answer-dependent style multiple-choice questions and demonstrate how the output allows for discrete analysis of questions based on various categories such as Fundamental Statements or Bloom's Taxonomy Levels. Additionally, we show how student demographics can be easily integrated to yield a holistic perspective on student performance in a course. The workflow offers dynamic grading opportunities for multiple-choice questions and versatility through its adaptability to assessment analyses. This approach to multiple-choice questions allows instructors to pinpoint factors affecting student performance and respond to changes to foster a healthy learning environment.

One of the inherent struggles of many 4-year public institutions relates to the high student-to-teaching team ratio (STR) (1). While much of the focus is on how STR directly affects student learning outcomes (2–4), there is the additional impact of grading hours that take away from contact hours. There are only so many teaching team hours available, which requires a balance of efficient and equitable grading practices with structured teaching time. If the goal is to encourage higher-order thinking in Bloom's Taxonomy (5–9), how can the instructor develop high-throughput summative assessments while also providing students with accurate and rapid feedback on their learning process?

The availability of grader hours can be heavily impacted when long-answer and short-answer questions are used in summative assessments. While these free-response questions are rigorous methods to assess student thought processes and application of information (10), they can quickly become overwhelming with an assessment entirely composed of free-response questions with limited graders and hundreds of students. Tools such as Gradescope have transformed the grading process as a platform tailored for scanned, pen-and-paper, and open-ended assessments and allow for detailed feedback by assessing submissions with a defined rubric (11). This software is constantly updating, and the benefits of this tool continue to grow.

Here, we describe an R-script we developed to adaptively grade multiple-choice questions (MCQ; R v4.2.1; dplyr v1.1.4). We can quickly grade, adjust for errors, and grade answer-dependent style questions. The data output of an MCQ assessment can then be easily aligned to Fundamental Statements (FSs) outlined by the American Society for Microbiology to help instructors link assessment to lesson plans and/or activity

performance (12, 13). Similarly, incorporating Bloom's can offer insights into cognitive complexity. Lastly, student demographic data can be incorporated to gain a deeper understanding of the diverse student academic outcomes. We envision this workflow as a powerful tool for instructors to maximize student outcomes and promote an inclusive learning environment in high-enrollment education.

PROCEDURE

The base script

The foundation of our script utilizes the "case_when()" command in dplyr (v1.1.4; Fig. 1). This script removes dependence on scantron grading software and provides users with a rapid and customizable grading scheme. We assume responses to MCQs are output with individual questions organized as columns, such as a scantron output. Using "case_when()," we can create our MCQ exam key and assign variable point values based on a single-answer choice or multiple-answer choices (Fig. 1A and B). We can also easily adjust scores when errors in the exam are found, such as during a regrade request.

Using this "case_when()" framework, we can build additional complexity in question type and immediately move to backend analysis to assess student learning and teaching strategies.

Answer-dependent multiple-choice questions

Answer-dependent MCQs provide students with the opportunity to choose their question progression, demonstrate their thought processes, and avoid an "all or nothing" question type. For the instructor, this is an opportunity to increase the complexity of MCQs by introducing free-response style questions in a sequential MCQ format (Fig. 1A). We describe two styles of answer-dependent questions, multi-step problem-solving and experimental design and assessment.

The most straightforward description of a multi-step MCQ would be a series of formulas. Students could be asked to use different formulas to calculate cell count, growth rate, and generation time in three separate questions based on the initial values provided. This could also be adapted to a series of reduction–oxidation reactions that follow the flow of electrons and calculate corresponding standard reduction potential, or through the tracking of nucleic acids in DNA replication and mRNA transcription eventually leading to translation and interpretation of a mutation (Supplemental Material).

From an experimental design and assessment perspective, students could be asked to reflect on various processes learned in class. For example, students could be asked to choose a microscope and sequentially match methods of implementation for a given specimen. Questions could ask students to explore challenges that might be encountered when using different methods of horizontal gene transfer. Lastly, data visualization interpretation could be broken down into steps of defining axes, exploring patterns, and describing levels of significance for a graph of the students' choosing.

All of these question types have been used thus far in an introductory microbiology course, and the implementation continues to grow.

Aggregation by question category

By breaking complex questions into simple questions, we can also further separate questions by question category, such as Bloom's, Fundamental Statements, Program Learning Objectives, or specific date (e.g., implementation of a new activity). This question category is inserted as a new row that can be used to aggregate outcomes and assess student learning and/or instructor teaching of a category of interest (Fig. 1C). This integration yields a holistic perspective on student performance, and instructors gain valuable insights into the multiple factors shaping student performance for their course (Fig. 1D and E).

A) Answer-dependent series

The following three questions are answer-dependent. Correctness of answers are dependent on previous answer.

7) Choose a macromolecule of your choice
A) Polysaccharides
B) Amino Acids
C) Lipids

8) Your macromolecule is closely associate with what food group?
A) Protein
B) Fats
C) Carbohydrates

9) Your food group represents the majority of what food item?
A) Olive Oil
B) Sugar
C) Egg White

B) Answer Key

```
df_qa <- df %>% mutate
    (P1 = case_when(Q1=="A" ~ 1),
     P2 = case_when(Q2=="A" ~ 1),
     P3 = case_when(Q3=="B" ~ 1),
     P4 = case_when(Q4=="C" ~ 1),
     P5 = case_when(Q5=="B" ~ 1),
     P6 = case_when(Q6=="D" ~ 1),
     P7 = case_when(Q7=="A" | Q7=="B" | Q7=="C" ~ 1),
     P8 = case_when(Q7=="A" & Q8=="C" |
                    Q7=="B" & Q8=="A" |
                    Q7=="C" & Q8=="B" ~ 1),
     P9 = case_when(Q8=="A" & Q9=="C" |
                    Q8=="B" & Q9=="A" |
                    Q8=="C" & Q9=="B" ~ 1))
```

C) df (Scantron Output)

SID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
12345	A	A	B	C	B	D	B	A	A
23456	A	A	A	B	A	B	C	D	A
34567	C	B	B	C	B	C	D	C	A
45678	A	C	B	C	A	D	A	A	D
56789	A	C	D	D	B	D	B	C	B
67890	B	C	B	A	B	C	A	A	A
Key	A	A	B	C	B	D	A	C	B
FS	1	1	2	2	3	3	2	1	3

D) df_a

SID	P1	P2	P3	P4	P5	P6	P7	P8	P9	total
12345	1	1	1	1	1	1		1		7
23456	1	1								2
34567			1	1	1					3
45678	1		1	1		1	1			5
56789	1				1	1			1	4
67890			1		1		1			3
Key	1	1	1	1	1	1	1	1	1	9
FS	1	1	2	2	3	3	2	1	3	18

SID	Demo
12345	coffee
23456	tea
34567	soda
45678	soda
56789	coffee
67890	tea

E) df_dist

Key	FS	Pcorrect	prop.A	prop.B	prop.C	prop.D	prop.E	mean
A	1	0.667	0.667	0.167	0.167	0	0	0.667
A	1	0.333	0.333	0.167	0.5	0	0	0.333
B	2	0.667	0.167	0.667	0	0.167	0	0.667
C	2	0.5	0.167	0.167	0.5	0.167	0	0.5
B	3	0.667	0.333	0.667	0	0	0	0.667
D	3	0.5	0	0.167	0.333	0.5	0	0.5
A	2	0.333	0.333	0.333	0.167	0.167	0	0.333
C	1	0.333	0.5	0	0.333	0.167	0	0.167
B	3	0.333	0.5	0.333	0	0.167	0	0.167

F) Graph by Theme & Demo

Demographic Analysis

one | two | three

FS Proportion Correct (y-axis: 0.00, 0.25, 0.50, 0.75, 1.00)

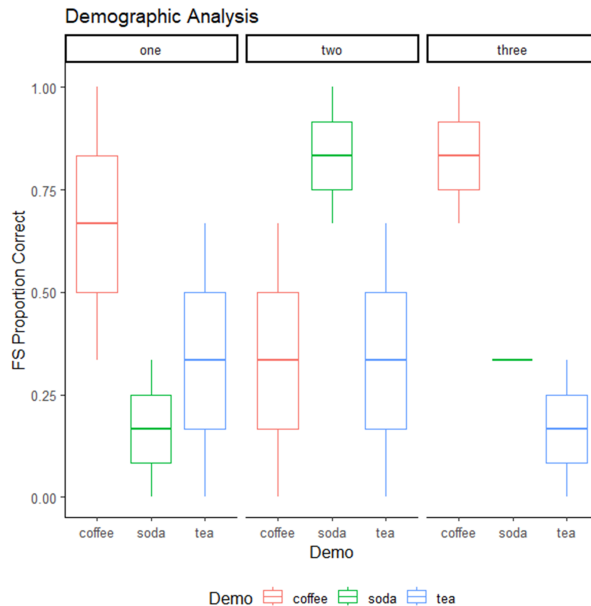Demo (x-axis: coffee, soda, tea)

Demo: coffee, soda, tea

FIG 1 A truncated flowchart for grading answer-dependent multiple-choice questions. An example series of three questions is provided (A) along with the corresponding syntax in R (B). A sample scantron output includes a question category row (C). The corresponding score output indicates individual question scores and total score (D). Data can be viewed in a tabular form to assess performance on individual questions (E) or go through an additional series of steps to graphically parse student performance by question category and/or student demographic (F). The full walkthrough is available in the Supplemental Material.

Disaggregation by student information

Disaggregation by student information may require access to and/or authorized use of data. Student information could include pre-confidence or time spent preparing (e.g., exam wrappers) (14). It could also include identity-based data. This data may be collected at the time of the summative assessment or accessed separately. In either case, these data are appended to a working data set to align with individual students' performance on the assessment. While this can provide a snapshot of where all students stand, coupling these data with question categories can be used to recognize disparities and opportunities for course-based interventions, whether that is an intervention for the students or the instructor themselves (Fig. 1F).

DISCUSSION

The high-throughput grading scheme described here can substantially reduce grading time, while balancing an equitable approach to grading with assessment questions at higher orders of thinking. The workflow we describe here is not meant to suggest a reduction in time spent on teaching, but a redirect of valuable teaching time toward increased student contact hours. This also provides additional flexibility when there are time or personnel constraints on grading, such as a reduction in teaching assistant availability or implementation of near-peer tutors who cannot serve as graders. We do not suggest this as a replacement to any assessment strategy, but merely as an additional approach to provide instructors with versatility in assessment design.

Answer-dependent MCQs provide an alternative method when the sheer number of students is an impediment to other assessment formats. The most straightforward approach is to incorporate answer-dependent MCQs into a summative assessment. The automated output of MCQs lends themselves well to post-analysis by the aggregation and disaggregation of data. The analysis could provide guided instructions on which FS a student needs to improve (e.g., through a FS focused retake) or provide the instructor with feedback to suggest improved study habits for lower-order Bloom's (e.g., studying vocabulary) or higher-order Bloom's (e.g., case–study analysis). Instructors could then integrate interactive activities that reinforce FS or provide real-world relevance of those FS. Moreover, making the concepts more relatable and helping students understand how to use the information in new situations will challenge students to engage at higher-order Bloom's.

While post-analysis is not limited to MCQs, the output may provide a more precise assessment of student performance on exams by definitively separating out individual questions by learning category (e.g., FS or Bloom's). This workflow provides instructors of high-enrollment courses with a high-throughput method to directly link higher-order assessment questions to a new or existing curriculum in an effort to rapidly respond to changes in student performance and foster a healthy learning environment.

AUTHOR AFFILIATIONS

[1]College of Biological Science, UC Davis, Davis, California, USA
[2]Department of Microbiology and Molecular Genetics, UC Davis, Davis, California, USA

AUTHOR ORCIDs

Kenjiro W. Quides http://orcid.org/0000-0003-2015-5264

AUTHOR CONTRIBUTIONS

Harnejan K. Atwal, Conceptualization, Software, Validation, Writing – original draft, Writing – review and editing | Kenjiro W. Quides, Conceptualization, Funding acquisition, Methodology, Software, Supervision, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Additional template material and code are available on GitHub (https://github.com/KWQevo/Answer-Dependent-Multiple-Choice-Questions).

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental Material (jmbe00047-24-s0001.docx). Annotated R Markdown file of code and example answer-dependent questions.

REFERENCES

1. Buckner E, Zhang Y. 2021. The quantity-quality tradeoff: a cross-national, longitudinal analysis of national student-faculty ratios in higher education. High Educ (Dordr) 82:39–60. https://doi.org/10.1007/s10734-020-00621-3

2. McDonald G. 2013. Does size matter? The impact of student-staff ratios. J Higher Educ Policy Manage 35:652–667. https://doi.org/10.1080/1360080X.2013.844668

3. Rujimora J, Campbell LO, DeMara RF. 2023. Exploring the student-to-faculty ratio and degree attainment in Florida. J Hispanic Higher Educ:153819272311725. https://doi.org/10.1177/15381927231172583

4. Barrasso AP, Spilios KE. 2021. A scoping review of literature assessing the impact of the learning assistant model. IJ STEM Ed 8. https://doi.org/10.1186/s40594-020-00267-8

5. Crowe A, Dirks C, Wenderoth MP. 2008. Biology in bloom: implementing Bloom's Taxonomy to enhance student learning in biology. CBE Life Sci Educ 7:368–381. https://doi.org/10.1187/cbe.08-05-0024

6. Anderson LW, Krathwohl DR. 2001. A taxonomy for learning, teaching, and assessing: a revision of Bloom's Taxonomy of educational objectives compete. Longman, New York.

7. Lemons PP, Lemons JD. 2013. Questions for assessing higher-order cognitive skills: It's not just Bloom's. CBE Life Sci Educ 12:47–58. https://doi.org/10.1187/cbe.12-03-0024

8. Stringer JK, Santen SA, Lee E, Rawls M, Bailey J, Richards A, Perera RA, Biskobing D. 2021. Examining Bloom's Taxonomy in multiple choice questions: students' approach to questions. Med Sci Educ 31:1311–1317. https://doi.org/10.1007/s40670-021-01305-y

9. Monrad SU, Bibler Zaidi NL, Grob KL, Kurtz JB, Tai AW, Hortsch M, Gruppen LD, Santen SA. 2021. What faculty write versus what students see? Perspectives on multiple-choice questions using Bloom's Taxonomy. Med Teach 43:575–582. https://doi.org/10.1080/0142159X.2021.1879376

10. Scouller K. 1998. The influence of assessment method on students' learning approaches: multiple choice question examination versus assignment essay. High Educ (Dordr):453–472.

11. Singh A, Karayev S, Gutowski K, Abbeel P. 2017. "Gradescope" L@S 2017; Cambridge Massachusetts USA, p 81–88ACM, New York, NY, USA. https://doi.org/10.1145/3051457.3051466

12. Briggs AG, Hughes LE, Brennan RE, Buchner J, Horak REA, Amburn DSK, McDonald AH, Primm TP, Smith AC, Stevens AM, Yung SB, Paustian TD. 2017. Concept inventory development reveals common student misconceptions about microbiology. J Microbiol Biol Educ 18:18.3.55. https://doi.org/10.1128/jmbe.v18i3.1319

13. Paustian TD, Briggs AG, Brennan RE, Boury N, Buchner J, Harris S, Horak REA, Hughes LE, Katz-Amburn DS, Massimelli MJ, McDonald AH, Primm TP, Smith AC, Stevens AM, Yung SB. 2017. Development validation, and application of the microbiology concept inventory. J Microbiol Biol Educ 18:18.3.49. https://doi.org/10.1128/jmbe.v18i3.1320

14. Hodges LC, Beall LC, Anderson EC, Carpenter TS, Cui L, Feeser E, Gierasch T, Nanes KM, Perks HM, Wagner C. 2020. National science teachers association effect of exam wrappers on student achievement in multiple, large STEM courses. J Coll Sci Teach 50:69–79.