Alternative grading: A more meaningful representation of student learning

•••

Ciaran Evans, Jessie Oehrlein, Sara Stoudt, Allison Thebold

Introductions

Session Outline

1. Warm-up
2. Philosophy of alternative grading
3. Some flavors of alternative grading methods
4. Shared experiences & reflections
5. Get your hands dirty!

Group Warm-up

(90-seconds)

Name one aspect of your current assessment / grading system you find unsatisfying.

(2-3 minutes)

Discuss with someone next to you

Armani ✔
@historyofarmani

what radicalized you?

9:06 PM · Aug 3, 2020 · Twitter for iPhone

18.1K Retweets and comments   **8.4K** Likes

What is "alternative" grading? 🤨

IN FEEDBACK LOOPS WE TRUST

CLEARLY DEFINED STANDARDS

HELPFUL FEEDBACK

MARKS INDICATE PROGRESS

REATTEMPTS WITHOUT PENALTY

from Grading for Growth,
Robert Talbert and David Clark (CC-BY-SA)

Standards-based Grading

Overview of Standards-Based Grading

- Structured mainly around a set of more granular learning goals ("standards" or "learning targets")
- Targets are re-assessable
- Final course grade mainly related to number of completed targets

Target 1	Target 2	Target 3
S	S	S
N	N	S
N	S	S
S	N	N
N	S	N
U	N	N
S	S	S
N	S	N
N	S	N
S	S	S
N	N	S
N	S	S
N	N	N
U	N	S
U	S	S
U	S	N
N	S	N
S	S	S

Defining Learning Targets

Most examples are pretty granular with many (15-25) targets for a term!

I preferred fewer targets but with clear sub-parts.

- Less difficult for students to track
- Sub-parts still provide clear expectations
- Atomizes understanding less
- **But** lower success on targets initially

1. **Identification of Key Information**
   - ☐ Identify population and variables in a given scenario.
   - ☐ Identify the variable type.
   - ☐ Distinguish between population and sample.
2. **Descriptive Statistics for Categorical Variables**
   - ☐ Calculate frequency and relative frequency.
   - ☐ Complete a contingency table.
   - ☐ Interpret and explain frequency/relative frequency tables in context.
   - ☐ Interpret and explain data visualizations for categorical variables in context.
3. **Descriptive Statistics for Numerical Variables**
   - ☐ Explain and interpret measures of center in context.
   - ☐ Explain and interpret measures of spread in context.
   - ☐ Identify summary statistics from a data visualization.
   - ☐ Describe, explain, and interpret data visualizations for numerical variables in context.

4. **Parameter Estimation**
   - ☐ Identify the parameter being estimated and calculate, explain, and interpret the point estimate in context.
   - ☐ Explain and interpret a confidence interval and margin of error in context.
   - ☐ Explain the role of confidence level, sample size, and sampling variability.
   - ☐ Draw and explain a conclusion based on a confidence interval in context.
5. **Testing Hypotheses**
   - ☐ Identify the parameter of interest and state null and alternative hypotheses for a given research question.
   - ☐ Explain and interpret a p-value in context.
   - ☐ Come to a conclusion based on a hypothesis test and explain the conclusion in context.
6. **Linear Regression and Correlation**
   - ☐ Identify response and explanatory variables.
   - ☐ Identify the sign and strength of a linear relationship based on the scatterplot or the correlation coefficient, and interpret that in context.
   - ☐ Interpret the parts of a linear regression model, including the coefficient of determination, in context.
   - ☐ Make predictions and draw conclusions based on a regression model and its limitations.

Assessing & Reassessing Learning Targets

- **Initial assessment:** For most folks using SBG, this is some kind of regularly-occurring quiz/test (but doesn't have to be!)
  - Online, untimed, out-of-class Checkpoint once every 3-ish weeks with two problems for each learning target we'd covered
  - Satisfactory/Not Yet Satisfactory (S/N) for attempted targets

- **Reassessment:** Often multiple options here!
  - Attempt again on new problems on the next Checkpoint
  - Write a reflection & revision on N problem and do one new written problem in office hours
  - Write a reflection & revision on N problem and do one new problem orally in office hours

Assigning a Final Grade

- Core of each grade level is learning target completion
- Can require multiple completions of some targets!
- Can bundle with other kinds of assignments (specifications-y)

	D-level (1.0)	**C-level (2.0)**	**B-level (3.0)**	**A-level (4.0)**
Learning targets	3 Satisfactory OR	3 Satisfactory	4 Satisfactory	6 Satisfactory
Project	Progressing	Progressing	Satisfactory	Excellent
Labs		3 Satisfactory	5 Satisfactory	7 Satisfactory
Reflections		Midterm or Final	Midterm and Final	Midterm and Final
Surveys & Homework		Regular completion	Regular completion	Regular completion

Specifications-based Grading

Overview of specs grading

Some core features:

- Assignments / problems graded on a binary ("mastered" vs. "not yet mastered") or other limited scale
- Clear specifications for what constitutes mastery
- Re-do attempts if an assignment is not mastered the first time
- Letter grades are earned by bundles of mastered assignments

Specifications for mastery

Example: GLMs, Project 1

Some sample requirements for mastering Project 1:

- All variables needed to address the research question are included in the EDA
- Appropriate visualizations are used to explore how each explanatory variable is related to the response. When needed, appropriate transformations are considered
- All plots shown in the report are explained and discussed in the text. What information does each graph provide, and why is that important for us to know when before building a model?

Specifications for mastery

Example: Intro Project

Some sample requirements for a Satisfactory project:

- Correct identification and explanation of errors in provided data visualizations.
- New visualizations are appropriate and correct the identified errors.
- Appropriate inference methods are used correctly, with any errors being minor and not related to understanding of statistical concepts.
- Conclusions are clearly stated in context and are supported by the analysis.

Grade bundles

Example: GLMs

To receive an **A**, students must:

- Master at least 5 HWs
- Master both projects
- Master both exams
- Master at least 2 challenge assignments

To receive a **B**, students must:

- Master at least 5 HWs
- Master one project
- Master both exams

Grade bundles

Example: Probability

To receive an **A**, students must:

- Master at least 10 HWs
- Earn an Excellent on final project
- Regularly present proofs / solutions in class
- Complete midterm and final reflections

To receive a **B**, students must:

- Master at least 8 HWs
- Earn a Satisfactory on final project
- Regularly present proofs / solutions in class
- Complete midterm or final reflection

Contract Grading

Opportunities to boost grade.

Core Contract - Everyone starts here.

If core not met, grade declines.

Overall Contract (+ details [here](#))

Assignment Type	Built-in Flexibility
Complete homework assignments on time (2-3 per week) with a good-faith effort	Can miss up to 4 without penalty
Complete lab assignments on time (0-1 per week)	One late, can revise in revision windows if incorrect
Complete each project (2 mini, 1 final) on time and meeting the conditions of "satisfactory" on the rubric	If "satisfactory" not met on mini-projects, meet with me to discuss improvements with no penalty

The Core - Contract for a B

Assignment	Purpose
Two extra peer reviews (classmates and/or MathStatBites)	Value: peer feedback and support + authentic audience
Writing Journal throughout semester (can take maximum two weeks off)	Value: frequent writing and reflection, requires planning to start early
Above and beyond work on one of three projects (make a written argument for why you think it meets the threshold)	Value: engagement and independence, provide evidence for argument
Turn a piece of work from the class into a piece for The Bucknellian (the student newspaper)	Value: revision for new context + authentic audience

Grade Boosters (B → B+, B+ → A-, A- → A)

Assignment Type	Consequence
Complete homework assignments on time (2-3 per week) with a good-faith effort	Miss between 5 and 8 → decline one step Miss between 9 and 12 → decline two steps…
Complete lab assignments on time (0-1 per week)	Two late/do not revise one → decline one step Three late/do not revise two → decline two steps…
Complete each project (2 mini, 1 final) on time and meeting the conditions of "satisfactory" on the rubric	Do not revise mini project or unsatisfactory final project → decline one step

When the core falters… (B → B-, B- → C+, etc.)

Ungrading

What is "ungrading"?

a pedagogical practice which entirely removes grades as a focus of the course

What is "ungrading"?

a pedagogical practice which entirely removes grades as a focus of the course

Instead...

- Exclusive focus on feedback, helping students accomplish the course learning objectives
- Marks on assignments are not attached to a grade outcome (e.g., A, B, C, D)
- Expectations for grades are <u>co-developed between students and the instructor</u>
- At the end of the course students make an argument for the grade they earned

Why no grades?

How do students argue their grade?

End of quarter portfolio with a selection of the work they completed:

- Two examples of code for each learning target
- Examples of extending thinking / challenging oneself
- Examples of revising thinking
- Examples of thoughtful collaborations

Learning Targets

The following learning targets speak to the big content ideas I'm designing this course for you to learn. I think these are probably exhaustive, and you will find that many of these targets overlap with each other.

Working with Data

- ❏ **WD-1:** I can import data from a variety[3] of formats (e.g., csv, xlsx, txt, etc.).
- ❏ **WD-2:** I can select necessary columns from a dataset.
- ❏ **WD-3:** I can filter rows from a dataframe for a variety[4] of data types (e.g., numeric, integer, character, factor, date).
- ❏ **WD-4:** I can modify existing variables and create new variables in a dataframe for a variety of data types (e.g., numeric, integer, character, factor, date).
- ❏ **WD-5:** I can use mutating joins to combine multiple dataframes.
- ❏ **WD-6:** I can use filtering joins to filter rows from a dataframe.
- ❏ **WD-7:** I can pivot dataframes from long to wide and visa versa

Reproducibility

- ❏ **R-1:** I can create professional looking, reproducible analyses using RStudio projects, Quarto documents, and the here package.
- ❏ **R-2:** I can write well documented and tidy code.
- ❏ **R-3:** I can write robust programs that are resistant to changes in inputs.

Data Visualization & Summarization

- ❏ **DVS-1:** I can create visualizations for a variety[5] of variable types (e.g., numeric, character, factor, date)
- ❏ **DVS-2:** I use plot modifications to make my visualization clear to the reader.
- ❏ **DVS-3:** I show creativity in my visualizations
- ❏ **DVS-4:** I can calculate numerical summaries of variables.
- ❏ **DVS-5:** I can find summaries of variables across multiple groups.
- ❏ **DVS-6:** I can create tables which make my summaries clear to the reader.
- ❏ **DVS-7:** I show creativity in my tables.

Program Efficiency

- ❏ **PE-1:** I can write concise code which does not repeat itself.
- ❏ **PE-2:** I can write functions to reduce repetition in my code.
- ❏ **PE-3:** I can use iteration to reduce repetition in my code.
- ❏ **PE-4:** I can use modern tools when carrying out my analysis.

Data Simulation & Modeling

- ❏ **DSM-1:** I can simulate data from a variety of probability models.
- ❏ **DSM-2:** I can fit a linear regression and extract necessary summary measures.

Similarities & Differences

How do you get students on board?

How do you mark assignments?

How are revisions handled?

How are deadlines handled?
___

Choose One Assignment

- Define standards:
  - What do you want students to learn from this?
- Think about how to give feedback:
  - What's common room for improvement?
  - How can you nudge students to improve in those ways?
- Decide what marks to give on it:
  - 3-level? 2-level? 1-level? Heck, 0-level?
- Decide how students can reattempt:
  - What opportunity do they have to care about your feedback?

Longer-term: ([Advice from Robert Talbert](#))

- Design backwards
  - Build your objectives (maybe like 20): What do you most care about?
  - What evidence might students give?
  - Design assessments to elicit that evidence
- Build a network (heyy!!)
  - A weird thing is going to go wrong in a way that nobody has contemplated before and you need someone to ask about it
- Keep it simple
  - You can iteratively refine as you go, so don't worry about "getting it right"

Thank you for joining us!

Biggest Success

Autumm Caines is a liminal space @Autumm · Jul 14

Any truth to the connection with grading meat?

💬 1   🔁   ♡ 3   ↥

Joshua Eyler @joshua_r_eyler · Jul 14

Yes, but in reverse. The meat industry developed its system after seeing it in action in our schools. And then they had conversations about how letters were not sufficient to capture the complexities of meat inspection, if you can believe it.

💬 3   🔁 1   ♡ 11   ↥

([link](#))