Developing a Playbook of Equitable Grading Practices

Stephen H. Edwards
edwards@cs.vt.edu
Computer Science
Virginia Tech
Blacksburg, VA, USA

David L. Largent
dllargent@bsu.edu
Computer Science
Ball State University
Muncie, IN, USA

Ben Schafer
ben.schafer@uni.edu
Computer Science
University of Northern Iowa
Cedar Falls, IA, USA

ABSTRACT

Traditional points-based grading can have negative impacts. As a result, many educators are experimenting with alternative grading practices that are more equitable for students. However, educators often face challenges in implementing equitable grading practices due to a lack of clear, practical descriptions of techniques and the fact that not all techniques are universally applicable. We propose a working group that will address this problem using a three-pronged strategy: conducting a systematic literature review to gather documented techniques, compiling "recipes" or concrete descriptions of these techniques, and publishing them in an open-source, online "playbook" of equitable grading practices as a community resource for educators. This approach aims to make such practices more accessible and adaptable to various classroom situations.

CCS CONCEPTS

• **Social and professional topics** → **Computer science education**; **Student assessment**.

KEYWORDS

Computer science education, Student assessment, Equitable grading, Specifications grading, Grading practices

1 INTRODUCTION AND GOALS

Grading student work is a core aspect of education. Traditional grading practices using points-based systems are ubiquitous, where numeric scores representing individual assignment grades are combined using variations of weighted averaging to determine an overall course grade. However, educators are more frequently recognizing that such approaches may have negative consequences [2, 8, 10]. As a result, educators are exploring alternative approaches to grading that aim to address drawbacks of traditional practices.

Such approaches are often labeled *equitable grading practices* (EGP), because they are less subject to the implicit biases embedded in traditional grading practices [2]. They also offer increased opportunities to accommodate factors outside the classroom that can negatively affect students, including work schedules, family life situations, or caretaker roles. These factors can disproportionately affect students from marginalized communities. By providing greater opportunities to reduce the grading impact of such external factors, which have nothing to do with content mastery or learning outcomes, such alternative practices can be more "equitable."

However, educators interested in EGPs often do not know where to start or how to adapt practices to their specific classroom situation. While many educators would explore alternative grading if they knew how, lack of easy access to clear descriptions of practical techniques is a significant obstacle. This situation is complicated by the fact that not all techniques work in every situation, instead embodying trade-offs that require consideration, with some being more advantageous for particular class sizes, or particular learning tools, or particular course content, etc. This working group will address this problem by compiling a readily available collection of accessible recipes for applying the most common equitable grading practices using a three-pronged strategy:

(1) Perform a literature survey to collect documented techniques and the source materials necessary for deeper study.
(2) Use experience reports uncovered in the literature survey together with experiences reported by group members to compile "recipes" or tactics grouped around thematic problem areas typically addressed by equitable grading practices.
(3) Publish an open-source, online *playbook* of these tactics using GitHub pages to provide a community resource for educators learning about equitable grading practices.

2 BACKGROUND

Traditional grading practices can reduce achievement, discourage students, and suppress effort [2, 8, 10]. Equitable grading practices aim to level the playing field for students of diverse backgrounds. While there are many approaches in use, there are common core aspects of these practices that recur in many grading approaches:

- A *reduced grading scale* that approximates pass/fail grading, eliminating partial credit. Often, a grading scale may have only 2-4 distinct levels, rather than using a 0–100 scale.
- Direct ties between the grading scale and the learning outcomes for each assignment, where "passing" the assignment means demonstrating the required learning outcomes.
- Avoidance of zero scores for missing work, instead encouraging students to revise and resubmit work that does not meet expectations until mastery has been achieved.

- Avoidance of using grades to reward or punish "behavior" (rather than learning) by separating concerns for time management, process, effort, etc., from the grading scale.

Feldman [2], Rapaport [7], and Nilson [5] advocate for reduced grading scales. These proposals address the disadvantages of 0–100 scales [2, 4, 5, 7, 8]. Nilson suggests a pass/fail grading scale, while Rapaport proposes a three-valued grading approach [7]. The EMRN (or EMRF) scale [9] uses a four-level grading scale.

EGPs emphasize handling late, missing, or insufficient work by allowing resubmissions, rather than assigning a zero score. This approach lowers the stakes, aligns with learning outcomes [5], and provides additional practice opportunities. Moreover, it supports reduced grading scales by enabling students to rework assignments until they demonstrate the required level of achievement, an approach that Bowen [1] argues is best for learning.

3 METHOD

As stated in the introduction, the **working group goals** are to **conduct a literature survey** that will be used to **compile recipes** of how EGPs are deployed and **publish an open-source, online playbook** of these recipes.

3.1 Phase 1: Literature Review

Using techniques for systematic reviews [6], the working group will define the scope of the search, use structured search techniques to identify candidate papers, and screen them for eligibility. The resulting references will be synthesized in Phase 2.

3.2 Phase 2: Compiling "Plays"

Results from the systematic review will be used to identify specific deployment strategies for individual EGPs. The working group will use the results of the systematic review to develop thematic categories for these practices. Examples might include:

- Use of reduced categorical/nominal grading scales.
- Flexible late policies.
- Approaches to allow resubmission/regrading of work that does not meet expectations.
- Approaches to determining final grades that avoid averaging or using zeroes for missing assignments.
- Rubric-based approaches to communicating the learning outcomes demonstrated in a successful assignment solution.
- Methods for encouraging good time management behavior separate from assignment grade penalties.
- Adapting traditional points-based online tools to work in EGP grading schemes.
- Encoding categorical/nominal grading scales into points-based gradebooks or learning management systems.

We expect the final categories to be determined by the results of the systematic literature review. These categories will be used to group practices into "chapters" in a playbook. The playbook will be organized as an online HTML book, implemented through a jekyll-based source repository hosted on GitHub and automatically rendered to HTML through GitHub's "pages" feature. All group members will have contributor access.

The systematic review will identify the "plays", and one group member will be assigned as the primary author for describing each play. Individual plays (or recipes) will be described in the spirit of design patterns [3], containing:

- A concise statement of the problem solved by the practice.
- A brief description of the tradeoffs involved, to clarify what classroom contexts are a good fit or a poor fit.
- A clear, operational description of the solution, or how the practice is deployed.
- A discussion section to capture any thoughts or community-driven feedback about the play.
- References citing the relevant publication(s).

3.3 Phase 3: Reviewing, Traceability, and Reporting

Each play in the playbook will be assigned two reviewers from the working group who will review, critique, and revise the section in coordination with its primary author. In addition, the working group will use a traceability matrix that maps each reference produced by the systematic literature review to its corresponding play(s). This will ensure that all literature review results are included where appropriate. The summary report will describe the process and highlight the most important elements of the playbook.

4 EXPECTED DELIVERABLES

The primary deliverables the working group will produce are the open-source, online playbook of EGPs. As discussed above, this will be hosted using a GitHub repository and delivered using GitHub pages, so that there will be a direct path for community-driven contributions, updates, and future evolution. In addition, a working group report will be produced that summarizes the process and highlights the contributions of the playbook.

ACKNOWLEDGMENTS

REFERENCES

[1] J. A. Bowen. 2012. *Teaching naked: How moving technology out of your college classroom will improve student learning.* John Wiley & Sons.
[2] J. Feldman. 2019. *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms.* Corwin Press.
[3] Eric Gamma, Richard Helm, Ralph Johnson, and John Vlissides. 1995. *Design Patterns: Elements of Reusable Software.* Addison-Wesley.
[4] Thomas R. Guskey. 2000. Grading Policies that Work Against Standards...and How to Fix Them. *NASSP Bulletin* 84, 620 (2000), 20–29. https://doi.org/10.1177/019263650008462003 arXiv:https://doi.org/10.1177/019263650008462003
[5] L. B. Nilson. 2015. *Specifications grading: Restoring rigor, motivating students, and saving faculty time.* Stylus Publishing, LLC.
[6] Matthew J. Page et al. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372 (2021). https://doi.org/10.1136/bmj.n71 arXiv:https://www.bmj.com/content/372/bmj.n71.full.pdf
[7] William J. Rapaport. 2011. A Triage Theory of Grading: The Good, the Bad, and the Middling. *Teaching Philosophy* 34, 4 (2011), 347–372. https://doi.org/10.5840/teachphil201134447
[8] D.B. Reeves. 2008. *Leading to Change / Effective Grading Practices.* Retrieved August 16, 2022 from https://www.ascd.org/el/articles/effective-grading-practices
[9] Robert Talbert. [n. d.]. *Giving marks that indicate progress.* Retrieved January 21, 2024 from https://gradingforgrowth.com/p/giving-marks-that-indicate-progress
[10] M. Townsley and T. Buckmiller. 2016. *What Does the Research Say about Standards-Based Grading? A Research Primer.* Retrieved August 16, 2022 from http://mctownsley.net/standards-based-grading-research/

ORGANIZATION PLAN

Member Selection Criteria

The primary criterion for member selection will be firsthand experience employing one or more alternative grading practices, which is instrumental in understanding the most important concerns to communicate to others who are considering employing a specific play. Absent such experience, a secondary criterion of experience with systematic literature reviews together with an awareness of existing literature on alternative grading practices is expected.

Collaboration Plan

Online work within the group will be loosely decentralized, with online meetings most frequently occurring among subgroups rather than the entire team. We will use Zoom or Microsoft Teams (depending on group member preferences) for online meetings, and will focus on the use of a github repository as the central shared workspace for developing the playbook and its resources. In addition, we will use Google drive or Microsoft Sharepoint for collaborating on meeting notes or sharing non-deliverable resources.

For coordinating meeting times, we will use timezone-aware meeting polling websites to determine appropriate time availability, and will rotate times to best accommodate all group members. We anticipate meetings involving the entire group once every two months, with subgroups consisting of members working on all of the plays within a single chapter to coordinate more regularly. Each chapter-based subgroup will have a specific group leader assigned to facilitate its coordination.

Significant portions of the work, including filtering of eligible literature identified during the literature review, writing of plays, and reviewing of plays, will be conducted in parallel and coordinated asynchronously using messaging on Google Chat or Microsoft Teams. Asynchronous messaging will help team members stay in regular contact throughout the project without requiring the overhead or timing constraints of frequent whole-team meetings.

Work Plan

As described in the proposal, the work plan consists of three phases to be completed before the conference, and a final fourth phase for final publication after the conference.

Phase 1: Literature Review. The first phase will be conducted in July and August, beginning with establishing the coordination mechanics among the group members. Major work during this phase will consist of identifying the scope of the literature review, the search query used to identify documents, and the selection criteria used to filter the search results. In parallel, the group will also establish the open-source license that will be used for the playbook (such as one of the Creative Commons licenses), and setting up of the basic starting point for the playbook repository on github.com. The final aspect of completing the literature review involves identifying the thematic categories arising from the literature review that will be used to organize the chapters of the playbook.

Phase 2: Compiling Plays. The second phase will be conducted in September and October and focuses on extracting and writing up the individual plays extracted from the literature. This begins with mapping the literature review papers to individual chapters and identifying the play(s) each contains. This information will be recorded in the beginning of a traceability matrix that will be finalized in Phase 3.

The group will use a basic bidding process to assign individual plays to the group member who will serve as the primary author for that play. Work on writing individual plays will be conducted primarily asynchronously with communication support through instant messaging. Writing during development will be stored directly in the playbook git repository on github.com, and all group members will be able to see the current status and progress of this work directly on the corresponding website.

Phase 3: Reviewing, Traceability, and Reporting. The third phase will be conducted in November. Writing the summary report will begin using an appropriate collaborative authoring tool such as Overleaf. A whole team meeting will be used to transition to reviewing tasks, where each group member will be assigned plays to review. The exact number of review assignments will depend on the number of plays extracted from the literature review, but currently we anticipate around 2 or 3 reviews assigned per group member.

The traceability matrix begun in Phase 2 will be completed and used to cross-check literature review articles against the plays in the playbook to ensure full coverage of the literature review results has been achieved. Phase 3 will conclude with completion of the working group summary report. In addition, the group expects to announce the playbook to the community through relevant mailing lists and solicit community contributions or additions at this time, as well as periodically every fall and spring going forward.

Phase 4: Final Publication. After the conference, the working group will collaborate together to refine the summary report for publication, coordinating efforts through direct messaging communications and a small number of online meetings as necessary.

Group Leaders

Stephen Edwards is Professor and Associate Department Head for Undergraduate Studies in the Department of Computer Science at Virginia Tech. He has employed multiple EGPs in his own teaching for three years and supported others adopting these techniques at his own institution. He has co-organized one ITiCSE working group and has participated in two others. He also has served in organizer roles for multiple conferences and workshops.

David Largent is Senior Lecturer in the Department of Computer Science at Ball State University. He has employed multiple EGPs in his own teaching (primarily specifications grading) since 2017 and supported others adopting these techniques at his own institution. He has co-organized six workshops and two Birds of a Feather sessions at SIGCSE TS, as well as numerous workshops and presentations at other conferences. He also has served in organizer roles for many local and regional conferences.

Ben Schafer is Professor of Computer Science and Program Coordinator for Computer Science Education at the University of Northern Iowa. He has employed EGPs in his classrooms since 2018 and has led initiatives to promote EGPs at UNI, in K-12 classrooms around the state of Iowa, and nationally. He has co-organized three workshops at SIGCSE TS and has been on the leadership team for multiple conferences and PD weeks.