Alternative Evaluation in CS Education Research: A Systematic Literature Map

Brian Harrington, Thezyrie Amarouche, Andrew Aucie, Shreeansha Bhattarai, Yuxin Chen, Raha Gharadaghi, Kaiyi Huang, Jenna Jiang, Linda Lo, Maliha Lodi, Fawaz Omidiya, Rohita Nalluri, Anagha Vadarevu

Department of Computer and Mathematical Sciences, University of Toronto Scarborough

UNIVERSITY OF TORONTO SCARBOROUGH

Abstract

There is a growing movement in the education community to re-evaluate grading practices and find ways of evaluating students that are more flexible, equitable, and focused on student growth. A wide variety of alternative evaluation systems have been developed, including mastery grading, specifications grading, contract grading and ungrading.

In this work, we perform a systematic literature search to find studies of alternative evaluation models in computer science education. A total of 37 papers were analyzed and coded on three metrics: Feature, Evaluation, and Population.

This map represents a tool for educators to find and evaluate research on alternative evaluation systems, resources for researchers to identify areas of future research, and a foundation for a future detailed literature analysis. The complete bibliography and literature map are available for download.

Search and Mapping Protocol

- Query: (”mastery grading” OR ”alternative grading” OR ”specification grading” OR ”specifications grading” OR ”mastery learning” OR ”specs grading” OR ”grading for equity” OR ”equitable grading” OR ”ungrading” OR “un-grading” OR ”contract grading”) AND (“programming” OR “computer science” OR “coding”)
- 376 papers retrieved from ACM Digital Library, Scopus, Science Direct and IEEE Xplore
- Inclusion Criteria: full papers, peer reviewed, measuring some form of alternative grading on some aspect of CS education - final corpus of 37 papers
- All papers double coded for criteria (Cohen’s Kappa = 0.7860), with discrepancies decided by group discussion

Dimensions and Categories

Intervention:

- Binary Grading
- Bucketed Grading
- Mastery Learning
- Re-submissions
- Minimum Grades
- Student Input

Evaluation Metric:

- Grades/Course Outcome
- Engagement
- Student Preferences
- Psychological Factors
- Teaching Workload
- Student Workload

Population:

- K-12
- Undergraduate
- Educators
- Introductory learners
- Gender
- Race
- Accessibility

Heat Map of Number of Papers by Category

Interventions vs Measurements

	Grades	Engage-ment	Prefer-ences	Psycho-logical Factors	Admin Workload	Student Workload
Binary Grading	8	3	8	2	5	6
Bucketed Grading	10	2	9	3	5	5
Mastery Learning	21	13	20	6	11	13
Resubmissions	15	10	18	3	8	14
Minimum Grades	1	1	1	0	0	1
Student input	2	0	2	1	1	2

Measurements vs Populations

	K-12	Under-graduate	Instruc-tors	Introduc-tory Course	Gender	Race	Accecssi-bility
Grades	1	27	2	9	2	2	2
Engagement	0	16	1	5	1	1	1
Preferences	1	27	3	7	1	1	1
Psychological	1	6	0	1	0	0	0
Admin Workload	0	14	4	4	1	1	0
Student Workload	0	18	2	4	2	2	1

Populations vs Interventions

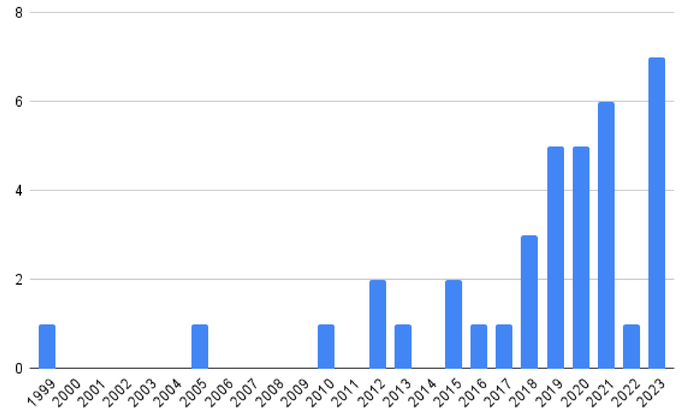
	Binary Grading	Bucketed Grading	Mastery Learning	Resub-missions	Minimum Grades	Student input on grade
K-12	0	0	1	1	0	0
Undergraduate	8	12	25	19	1	2
Instructors	1	2	2	3	0	0
Introductory Course	2	3	11	5	0	0
Gender	1	1	2	2	0	0
Race	1	1	2	2	0	0
Accecssibility	1	0	1	1	1	0

Resources & Downloadable Map

The complete mapping is available as either a PDF or downloadable CSV with accompanying bibliography file at

https://github.com/BrianHarringtonUTSC/SIGCSE2024AltGradingLitMap

Paper Counts by Year



Findings & Future Work

- Rapidly growing area: over 50% of papers published in past 5 years)
- Very little consistency in terminology and intervention selection across studies
- Inconsistent implementation of specific interventions: e.g., “unlimited” re-submissions or number of “buckets” in bucketed grading
- Imbalance in study density: many studies on undergraduate + grades or mastery learning + preferences, but only 4 papers evaluating impact on educators, and only 2 analysing by race/gender

References

[1] Clark, D., and Talbert, R.
Grading for Growth: A Guide to Alternative Grading Practices that Promote Authentic Learning and Student Engagement in Higher Education.
Taylor & Francis, 2023.

[2] Feldman, J.
Grading for equity: What it is, why it matters, and how it can transform schools and classrooms.
Corwin Press, 2018.

[3] Harrington, B., Galal, A., Nalluri, R., Nasiha, F., and Vadarevu, A.
Specifications and contract grading in computer science education.
In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education* (2024), SIGCSE 2024.

[4] Kitchenham, B., and Charters, S. M.
Guidelines for performing systematic literature reviews in software engineering.
Tech. rep., Technical report, ver. 2.3 ebse technical report. ebse, 2007.