Specifications and Contract Grading in Computer Science Education

Brian Harrington, Abdalaziz Galal, Rohita Nalluri, Faiza Nasiha and Anagha Vadarevu

Department of Computer and Mathematical Sciences, University of Toronto Scarborough

Abstract

With the recent growth in popularity of alternative evaluation methods, two methodologies have become particularly prevalent in CS education literature: Specifications grading and contract grading. Recent work has shown that these novel evaluation approaches can have positive impacts in the classroom and lead to more equitable outcomes for students. However, there is not yet a consensus on terminology, implementation details and best practices. In this work, we review the literature on the use of specifications and contract grading in CS education. We find that while there is a good deal of promising research, there is a great deal of variation in methodologies and a sparsity of evaluation of the efficacy of learning outcomes.

Methodology

- Query: **`"specification grading" OR "specifications grading" OR "specs grading" OR "contract grading"`**
- Search run on ACM Digital Library, Scopus, Science Direct and IEEE Xplore. 41 papers retrieved, 10 meeting inclusion criteria
- Double coded for inter-rater reliability: Fleiss-Davies kappa = 0.75, discrepancies decided by mediation.
- Snowballing process resulted in 10 additional papers, 1 meeting inclusion criteria

Research Questions

1. Where has specifications or contract grading been implemented and evaluated in CS education research?
2. How widely have implementation details varied?
3. How have the implementations been evaluated?
4. Why have specifications or contract grading been implemented? What motivations and justifications have been identified by the authors?
5. What obstacles have been identified for the implementation of specifications and contract grading?
6. What benefits have been demonstrated by implementing specifications or contract grading?
7. What gaps can be identified in the research that require future exploration?

References

[1] J. Aycock and J. Uhl. Choice in the classroom. *SIGCSE Bull.*, 37(4):84–88, dec 2005. ISSN 0097-8418.

[2] E. Dosmar and J. Williams. Student reflections on learning as the basis for course grades. In *2022 ASEE Annual Conference & Exposition*, 2022.

[3] P. Gestwicki. Design and evaluation of an undergraduate course on software development practices. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, SIGCSE '18, page 221–226, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450351034.

[4] P. Gestwicki. Godot engine and checklist-based specifications: Revising a game programming class for asynchronous online teaching. *J. Comput. Sci. Coll.*, 37(4):30–40, oct 2021. ISSN 1937-4771.

[5] N. LeJeune. Contract grading with mastery learning in cs 1. *J. Comput. Sci. Coll.*, 26(2):149–156, dec 2010. ISSN 1937-4771.

[6] G. M. Mirsky. Effectiveness of specifications grading in teaching technical writing to computer science students. *J. Comput. Sci. Coll.*, 34(1): 104–110, oct 2018. ISSN 1937-4771.

[7] K. R. Sanft, B. Drawert, and A. Whitley. Modified specifications grading in computer science: Preliminary assessment and experience across five undergraduate courses. *J. Comput. Sci. Coll.*, 36(5):34–46, jan 2021. ISSN 1937-4771.

[8] S. Spurlock. Improving student motivation by ungrading. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, SIGCSE 2023, page 631–637, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394314.

[9] E. Tuson and T. Hickey. Mastery learning and specs grading in discrete math. In *Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education Vol. 1*, ITiCSE '22, page 19–25, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450392013.

[10] E. Tuson and T. Hickey. Mastery learning with specs grading for programming courses. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, SIGCSE 2023, page 1049–1054, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394314.

[11] R. Weber. Using alternative grading in a non-major algorithms course. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, SIGCSE 2023, page 638–644, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394314.

Variations & Motivations

Authors	Date	Course	Variations			
			Binary Grading	Bucketed Grading	Re-submissions	Standards Based Evaluation
Aycock and Uhl [1]	2005	undergraduate CS				✓
LeJeune [5]	2010	CS1	✓		✓	✓
Gestwicki (2018) [3]	2018	undergraduate CS		✓		✓
Mirsky [6]	2018	undergraduate CS	✓		✓	✓
Gestwicki (2021) [4]	2021	undergraduate CS		✓		✓
Sanft et al. [7]	2021	CS1 + undergraduate CS	✓		✓	✓
Dosmar and Williams [2]	2022	CS1			✓	✓
Tuson and Hickey (2022) [9]	2022	undergraduate discrete math	✓		✓	✓
Spurlock [8]	2023	undergraduate CS		✓	✓	✓
Tuson and Hickey (2023) [10]	2023	undergraduate CS	✓		✓	✓
Weber [11]	2023	undergraduate and graduate non CS-majors		✓	✓	✓

Table 1:Summary of Variations on Contract/Specifications Grading by Paper

	Motivations				Obstacles		
	Diverse Learning Cohorts	Guaranteeing Mastery	Student Control	Instructor Workload	Procrastination	Scaling Delivery	Instructor Workload
Aycock and Uhl [1]			✓	✓		✓	✓
LeJeune [5]		✓					✓
Gestwicki (2018) [3]							✓
Mirsky [6]		✓		✓			
Gestwicki (2021) [4]					✓		
Sanft et al. [7]		✓		✓			
Dosmar and Williams [2]			✓				
Tuson and Hickey (2022) [9]	✓			✓	✓		✓
Spurlock [8]			✓		✓		✓
Tuson and Hickey (2023) [10]		✓		✓	✓	✓	
Weber [11]		✓		✓	✓		

Table 2:Summary of Stated Motivations for Adopting Alternative Grading and Obstacles Encountered

Findings

- All research was done in North American institutions with most studies focused on undergraduate level CS
- Papers had variation in implementation details:
  - Categorical Evaluation: Variations in cutoff for binary grading. Bucketed grading had variations in number, level and definition of buckets
  - Re-submission Limitations: Limitations exist in form of either time-window, set number of attempts, or penalty/decaying grade policies
  - Standards Based Grading: No standard or best practice for translating components to final course grades
- Variety of motivations for adopting grading methodologies
  - Most common: reduction in instructor workload for grading
- Different obstacles to adoption mentioned
  - Most common: student procrastination and initialization increase in instructor workload
- Evaluation was based mostly on student preferences and perceived efficacy
  - General findings were positive with respect to student evaluation, perceived stress, and self assessment of learning
- No attempts to evaluate learning outcomes or retention compared to traditional grading systems