Short vs. Extended Answer Questions in Computer Science Exams

Alejandro Salinger

Opportunities and New Directions

April 26th, 2012

ajsalinger@uwaterloo.ca

Computer Science Written Exams

- Many choices of question formats
  - Multiple-choice
  - True/false
  - Short answer
  - Problem solving
  - Code writing

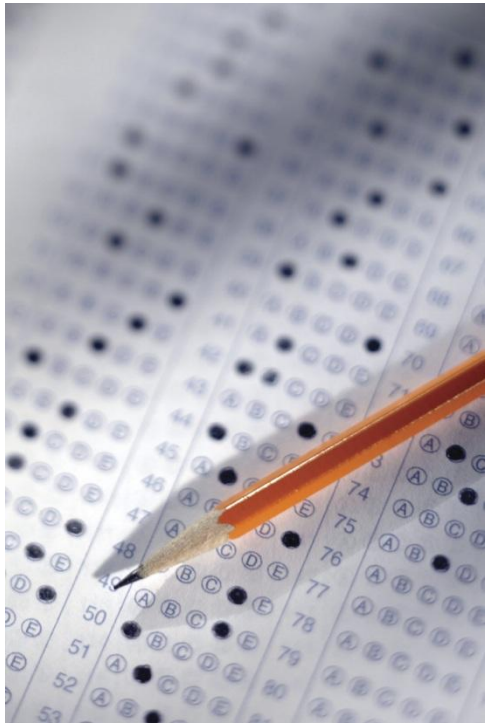
- How suitable is each type for CS exams?
- What are their implications?

Outline

- Short and Extended Answer Questions
- Do both measure the same skills?
- Instructors' perspectives
- Influence on learning
- Discussion
  - Intended Learning Outcomes
  - Assessments as a learning instance
  - Structural fidelity

Short Answer Questions

- Multiple-choice
- True-or-false
- Fill-in-the-blank
- Brief-answer



What are the best-, average-, and worst-case times to sort n items using Quicksort?

Extended Answer Questions

- Code Writing
- Problem Solving
- Mathematical Proof

Prove that in the comparison model any sorting algorithm requires $\Omega(n \log n)$ comparisons in the worst case when sorting n items.

Pros and Cons

Short Answer Questions

- Efficient administration
- Objective grading
- Timely feedback
- Can test wide range of topics
- Independent of writing skills
- Easy to evaluate test itself

Extended Answer Questions

- Take less time to construct
- Easier to test high levels of learning
- Partial credit
- No guessing
- Test writing abilities
- Higher structural fidelity

Can both format types measure the same skills?

Can both formats measure the same?

- Can multiple-choice replace constructed-response?
- CR items provide less information in more time and at greater cost than MC (Lukhele, Thissen, Wainer, 1994)
- Score of essay adds minimal information about grade beyond MC score (Wastad & Becker, 1994)
- Little support for stereotype of MC and CR measuring different constructs (Bennet, Rock, Wang, 1991)

Can both formats measure the same?

- Some skills are too complex to be measured effectively with MC questions
- Measures can change over time

(Livingston, 2009)

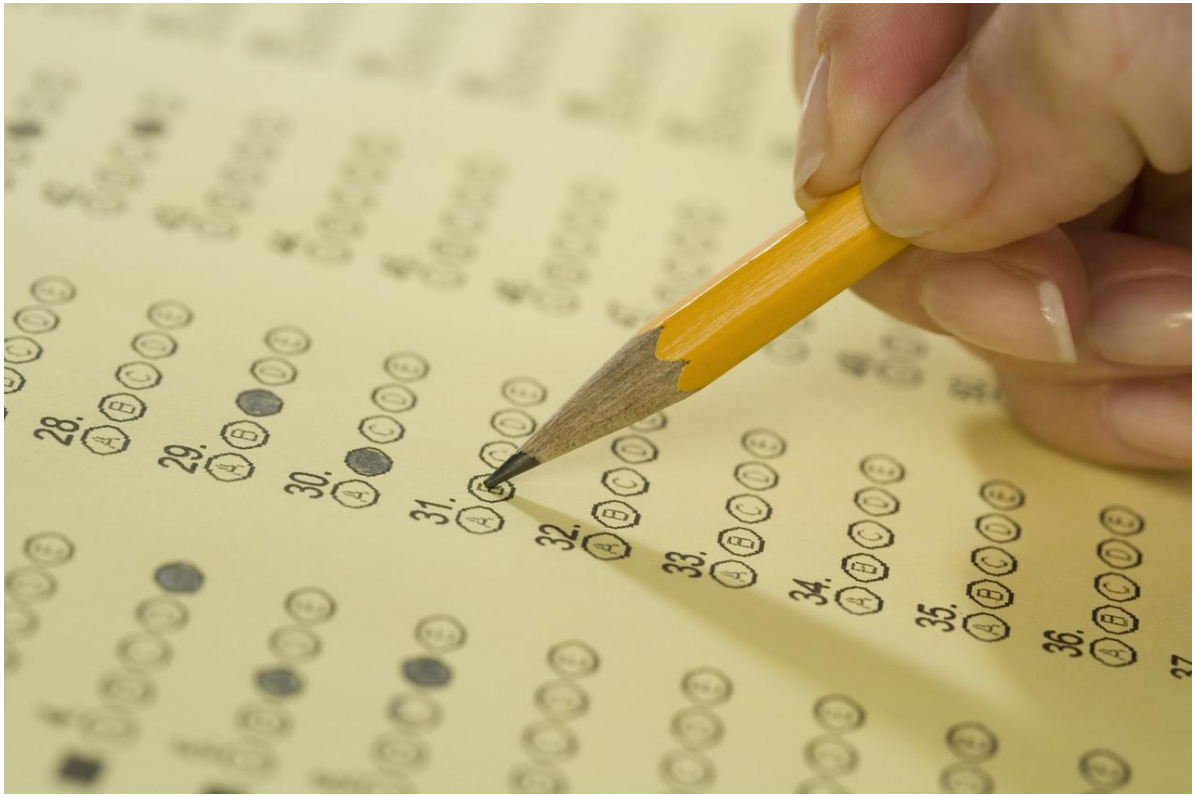
- MC questions are not homogeneous
  - Substantial differences in MC questions
  - Extremely difficult to construct MC question at Application level
  - No combination of MC questions exactly matches skills of some CR questions
  - Weak correlations at same difficulty level

(Simkin, Kuechler, 2005, 2010)

- The evidence is inconclusive

(Martinez, 1999)

Computer Science instructors perspectives on multiple-choice questions



Why do you use MC questions?

(Shuhidan, Hamilton, D'Souza, 2010)

- Understanding
  - "To get an idea of the breadth of students' understanding"
- Determine level of knowledge
  - "MC can also be used to test the depth of knowledge"
- Confidence
  - "To give weaker students confidence to answer questions"
- Shorter feedback time
- Easy question
  - "Keep the weaker students on track"
  - "To test understanding of fundamental terms/phrases used in programming"
- Student-centered
  - "To keep students happy"
  - "To constrain the students' creativity"
  - "To prepare students for later courses where this will be even more common"

CS Instructors perspectives on MC

10% do not support the use of MC:

- "I have NEVER used multiple choice questions in an exam!"
- "You need to include essay questions because Computer Science students need to know how to write"
- "I feel that multiple choices is a completely inappropriate tool for judging deep understanding and comprehension of programming concepts"

Influence on learning approaches

Influence on learning approaches

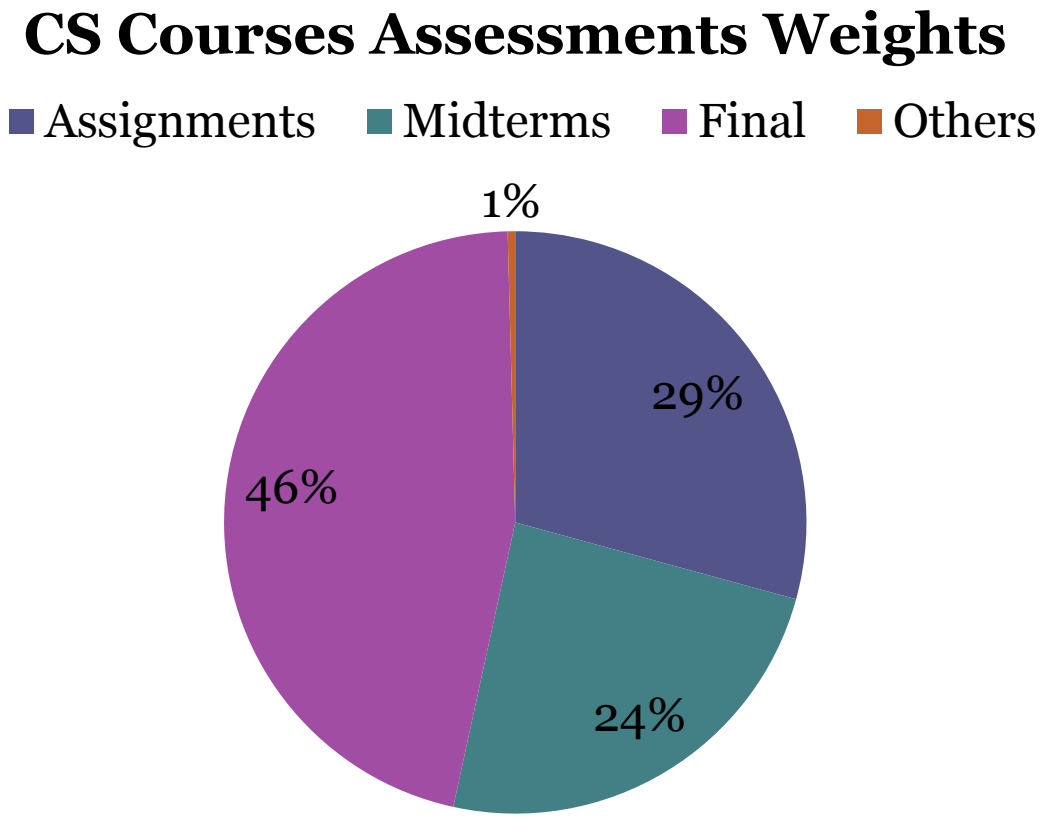
- Multiple-choice:
  - Surface approach
  - Knowledge-based skills
- Essay
  - Deep approach
  - Comprehension, application, analysis
- Preference:
  - Essays → deep approach and better performance in essays
  - MCQ → surface approach and worse performance in essays
- Deep approach and perception of MCQ assessing higher levels → poor performance

(Scouller, 1998)

Discussion

Intended Learning Outcomes

- ILOs go beyond Knowledge and Comprehension in most CS courses
- Difficult to measure these with short-answer questions
- If tests rely on short-answer questions,
  - High-level skills not measured
  - Likely not attained
- What about other assessments?


CS Courses Assessments Weights
Assignments
Midterms
Final
Others
1%
29%
46%
24%

Assessment as a Learning Instance

Show how to sort n integers in the range $[0, n^2 - 1]$ in $O(n)$ time.

- Exam with a few long questions enables:
  - Reflecting
  - Deriving ideas
  - Making connections
  - Evaluation of ideas
  - Creativity
  - Writing skills
- Valuable feedback for instructors and students

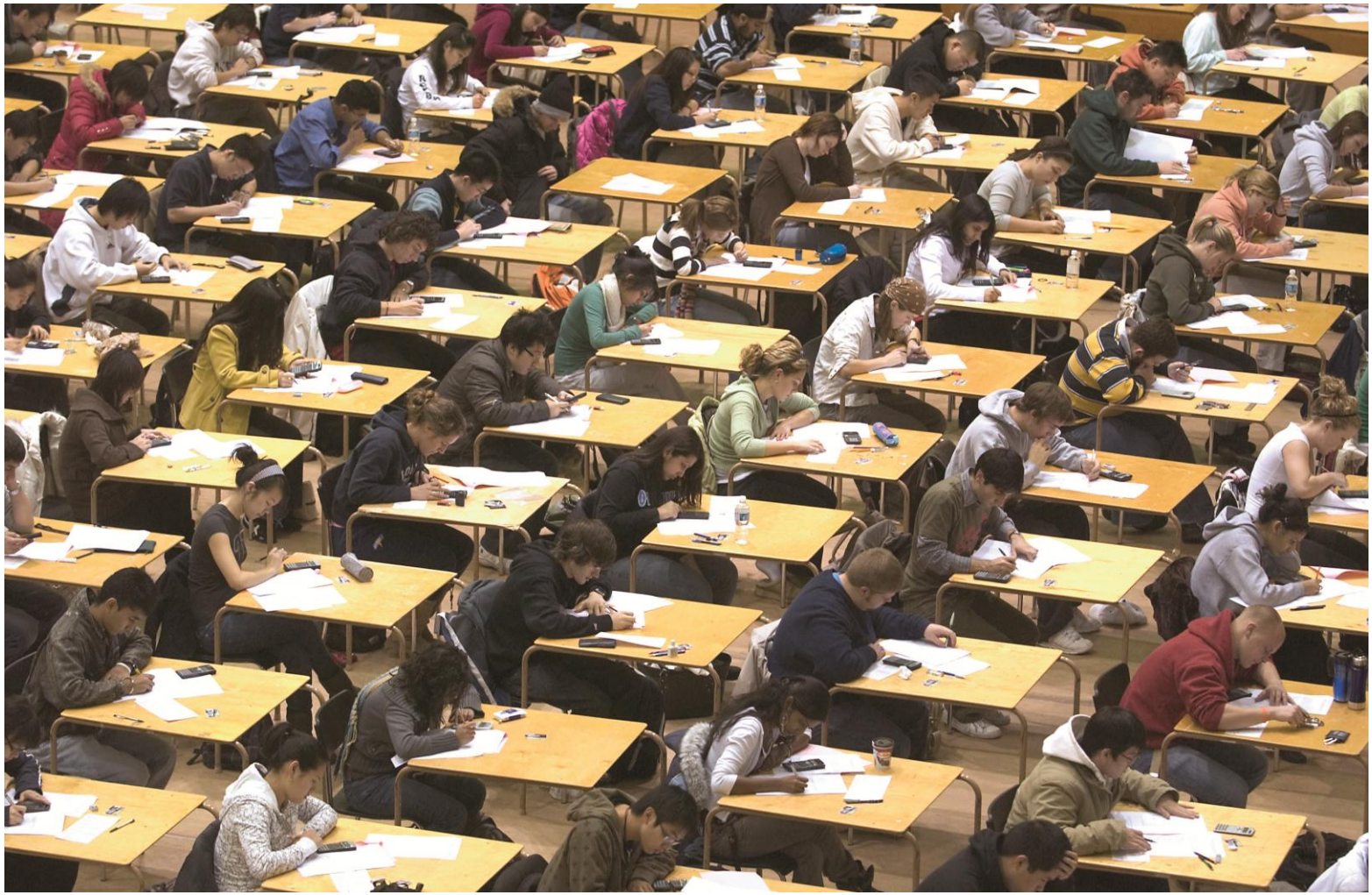
Structural Fidelity

- Easier to achieve with extended-answer questions

- For example, computer programming:
  - Writing a program closer to real-world situation
  - Skills measured by short-answer questions not very useful
  - Extended-answer questions can be complemented by provided aid

Conclusions

- Choice of question format has an important influence on learning assessment, learning approaches, and perceptions
- Skills involved in Computer Science courses are better measured with extended-answer questions
- Extended-answer questions should be used in CS courses from the first year, with the appropriate weight in the final grade



Thank you

References

- CUT project: http://www.cs.uwaterloo.ca/~ajsalinger/cut_project.pdf
- Bennett, R. E., Rock, D. A., & Wang, M. (1991). Equivalence of free-response and multiple-choice items. *Journal of Educational Measurement* , 28(1), pp. 77-92.
- Kuechler, W. L., & Simkin, M. G. (2010). Why is performance on multiple-choice tests and constructed-response tests not more closely related? theory and an empirical test*. *Decision Sciences Journal of Innovative Education*, 8(1), 55–73.
- Livingston, S. A. (2009). Constructed-response test questions: Why we use them; how we score them. ETS, R&D Connections 11.
- Lukhele, R., Thissen, D., & Wainer, H. (1994). On the relative value of multiple-choice, constructed response, and examinee-selected items on two achievement tests. *Journal of Educational Measurement* , 31(3), pp. 234-250.
- Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35, 453-472.
- Shuhidan, S., Hamilton, M., & D'Souza, D. (2010). Instructor perspectives of multiple-choice questions in summative assessment for novice programmers. *Computer Science Education*, 20(3), 229-259.
- Simkin, M. G., & Kuechler, W. L. (2005). Multiple-choice tests and student understanding: What is the connection? *Decision Sciences Journal of Innovative Education*, 3(1), 73–98.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education*, 30(4), 325-341.
- Walstad, W. B., & Becker, W. E. (1994, May). Achievement differences on multiple-choice and essay tests in economics. *American Economic Review*, 84(2), 193-96.