COMPUTERISED ADAPTIVE TESTING
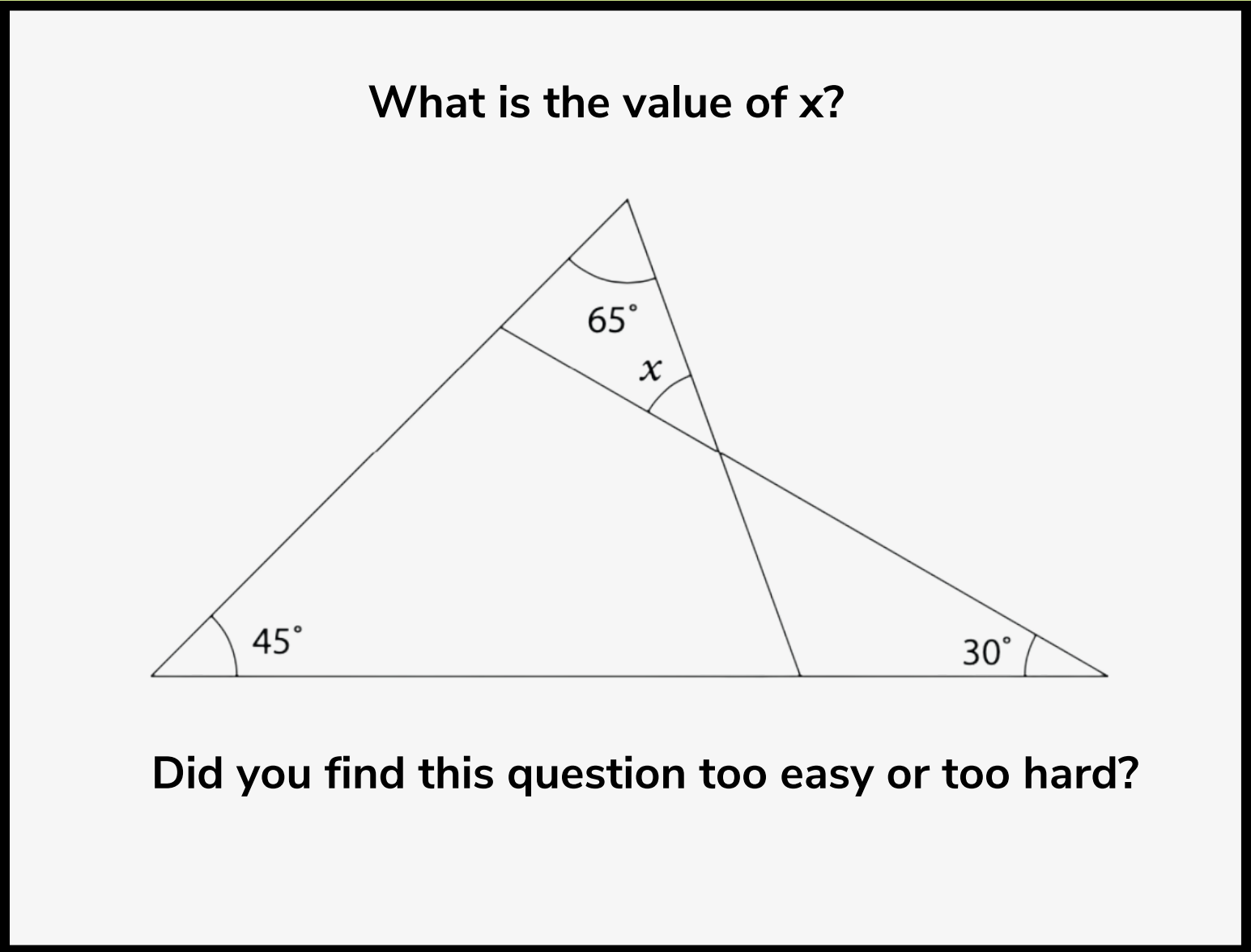
PROJECT 52

STUDENTS Hajin Kim
Oscar Li

SUPERVISORS Yu-Cheng Tu
Andrew Meads

INTRODUCTION

Computer adaptive testing (CAT) is a novel testing paradigm that adjusts question difficulty based on the user's previous responses.

Multistage testing (MST) is a form of CAT that separates a test into question sets (testlets) of different difficulties. By having multiple stages, an algorithm can be used to select the testlet that is most appropriate for a particular user at each stage.

Using MST, we developed a testing platform which aims to provide more accurate ability estimates while reducing the overall test length compared to traditional fixed length tests.

What is the value of x?

65°
x
45°
30°

Did you find this question too easy or too hard?

WHAT WE MADE

For our project, we have implemented an online platform that administers multistage and fixed-length tests targeting high school mathematics students.
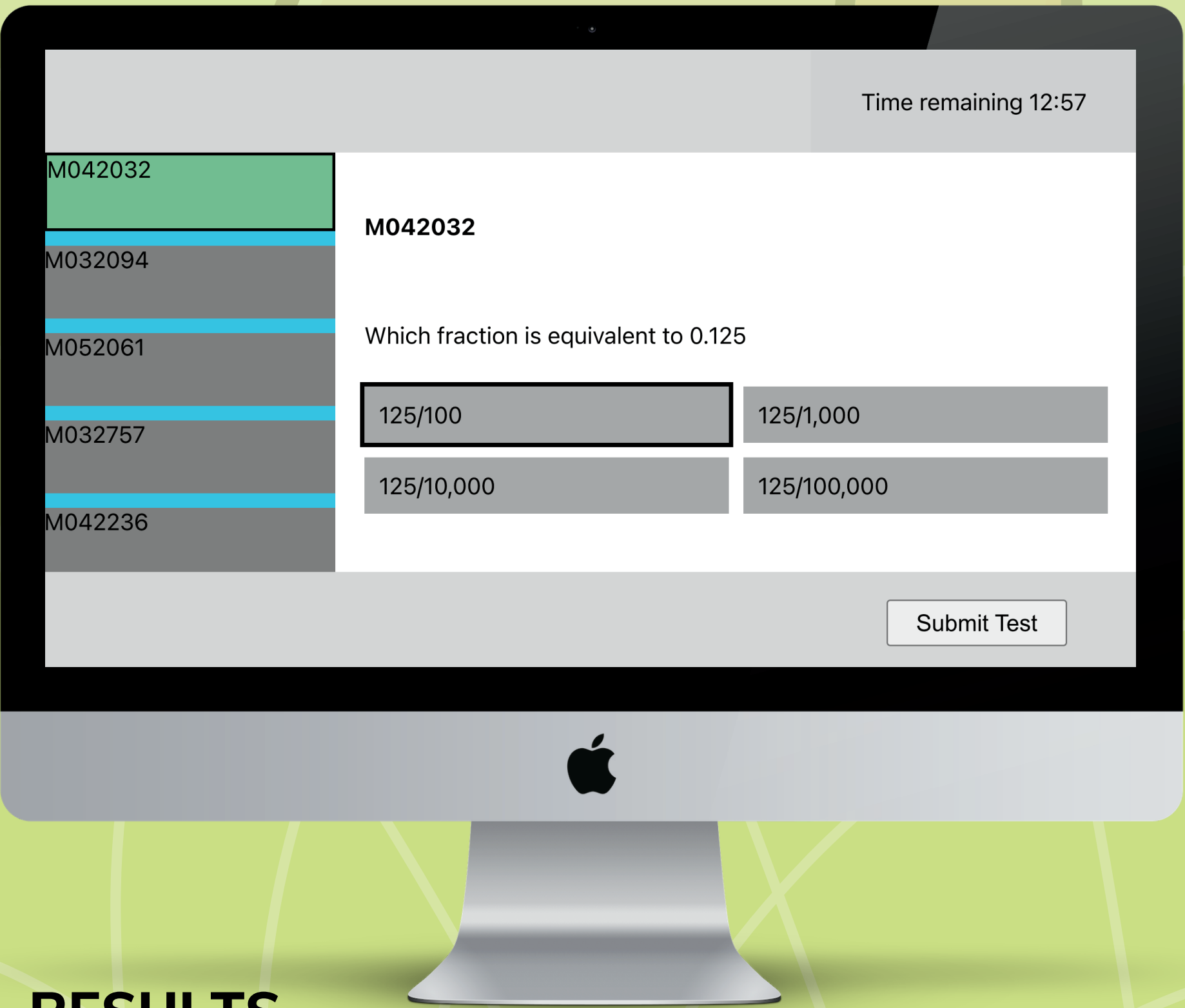
Features include:

- Support for both short answer questions and MCQ
- Leverages implementations of mstR package [1]
- Standard features of a testing platform, such as a timer, login, and score persistence.
- Developer API for creating and deleting questions and users

FEEDBACK

"I preferred the fixed length test over the multistage test because I found that I can focus more if I do all the questions in one go rather than having a rest in between."

"The UI is clean, no unnecessary information is provided. It's easy to see which question is completed, and the instructions are relatively clear."

Time remaining 12:57

M042032
M032094
M052061
M032757
M042236

M042032

Which fraction is equivalent to 0.125

125/100
125/1,000
125/10,000
125/100,000

Submit Test

RESULTS

Preliminary testing on peers showed that 40% of participants found multistage testing less tedious than fixed-length tests and 60% found that the questions in the multistage test were better suited to their ability.
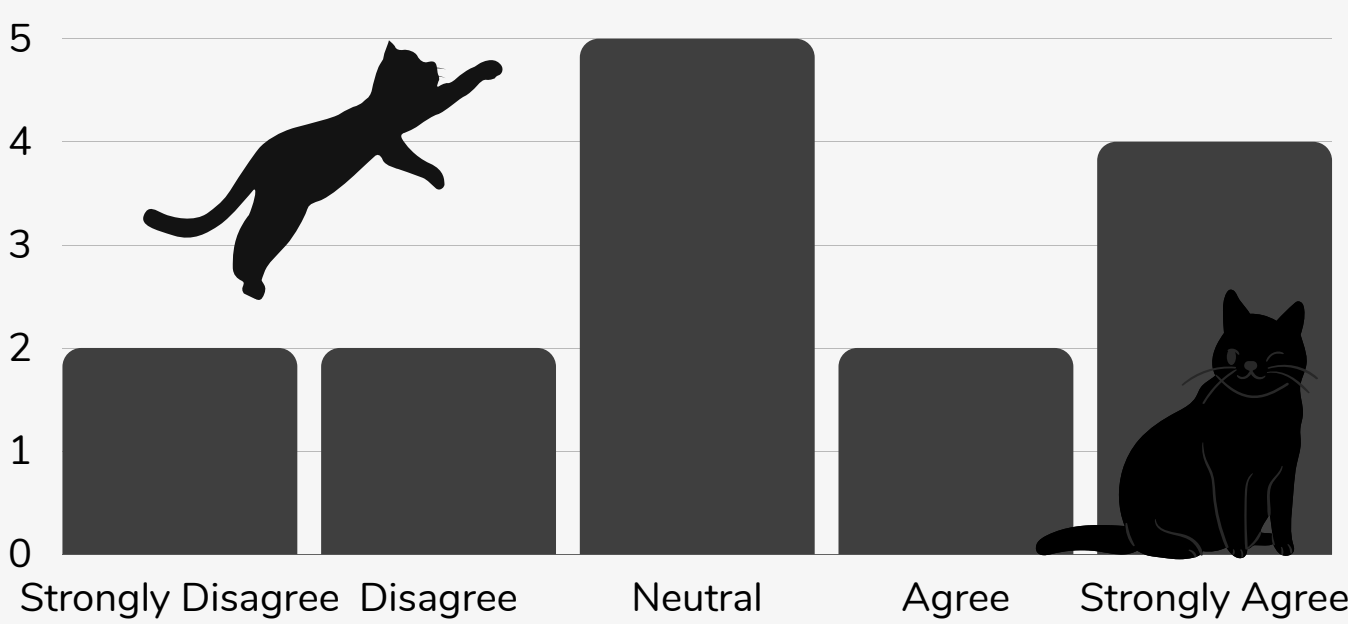
Post-Test Survey: I found that the multistage test was less tedious to get through than the fixed length test.
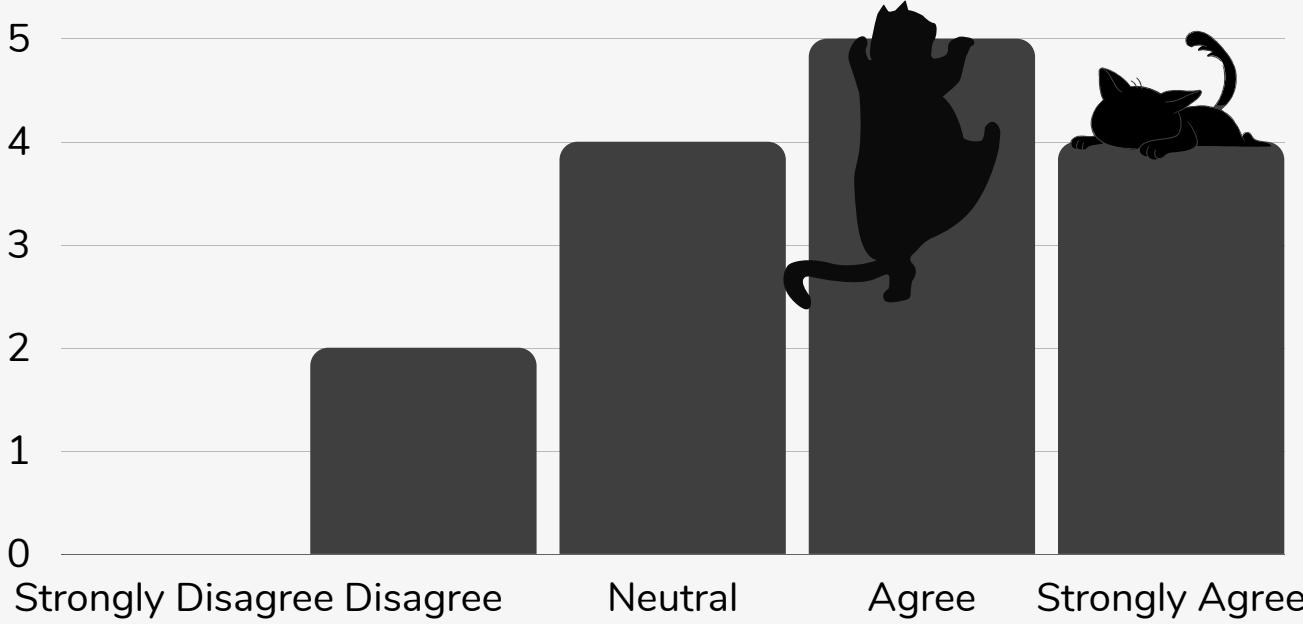
Post-Test Survey: The multistage test questions were better suited to my ability than the fixed length test

CONCLUSION

- Developed a tool that allows researchers to compare MST experimentally against the fixed-length test.
- The post-test survey from pilot testing shows a promising prospect for the new testing paradigm.
- Further fine-tuning of the platform is needed to reliably claim the feasibility of MST in various settings.

FUTURE WORK

- Acquire ethics approval to conduct a formal study on high school students
- Fine-tuning parameters used for item bank generation, score estimation, and testlet administration in MST
- Implement custom input types for certain questions.
- Improve user interface for the target audience

[1] Package mstR, 2018. https://cran.r-project.org/web/packages/mstR/mstR.pdf