To Test or Not to Test?

***

An Intelligent Recommender System for Personalized Test Scheduling

Okan Bulut
University of Alberta

Okan Bulut

Damien Cormier

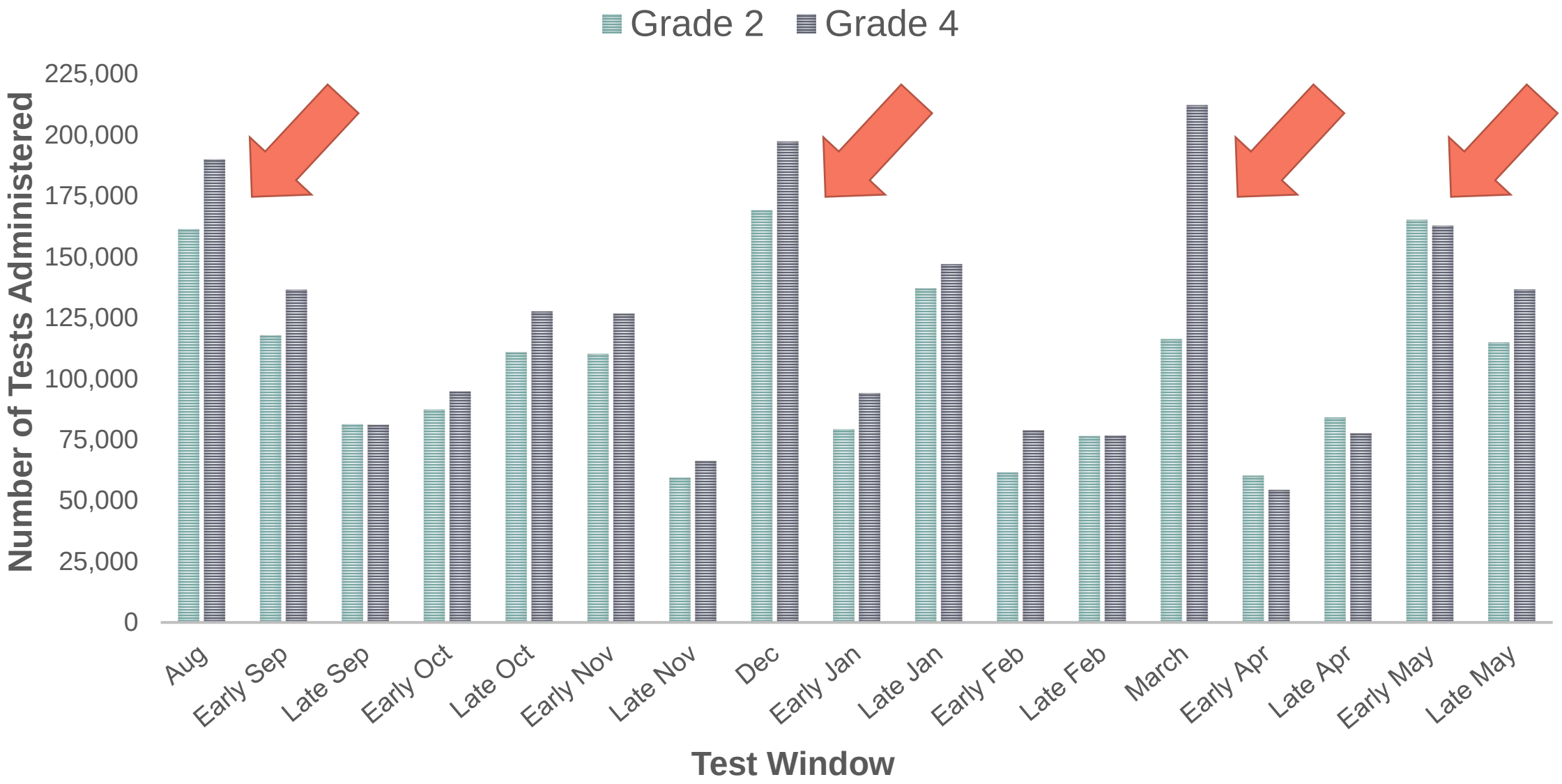
RENAISSANCE®

Jinnie Shin

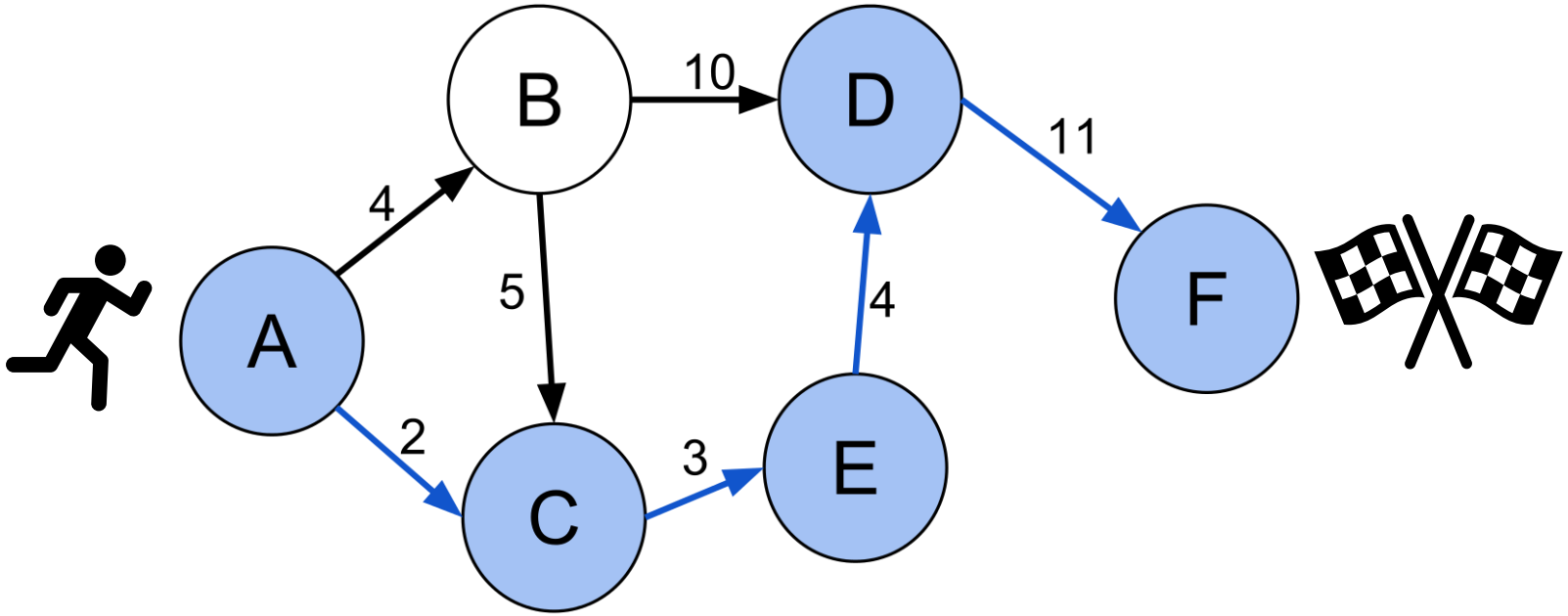
STAR™ Math

Grade 2: 355,078 students

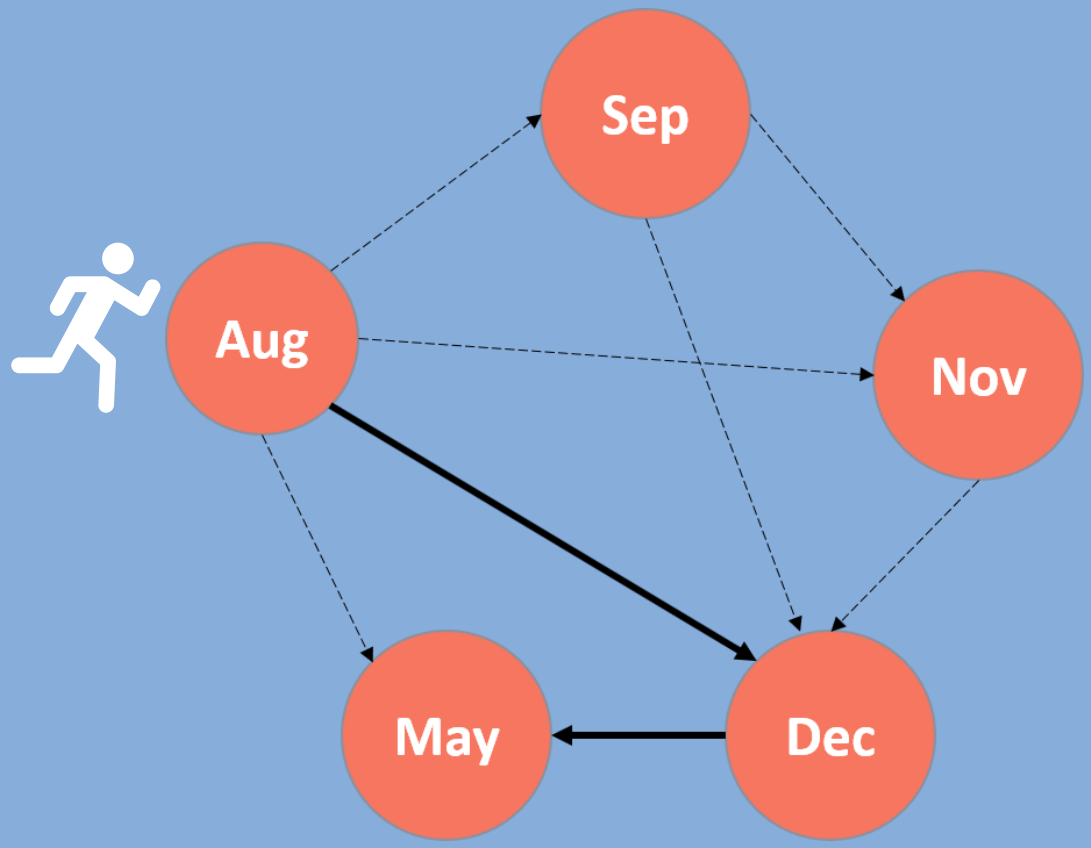
Grade 4: 390,336 students


Grade 2
Grade 4
Number of Tests Administered
225,000
200,000
175,000
150,000
125,000
100,000
75,000
50,000
25,000
0
Aug
Early Sep
Late Sep
Early Oct
Late Oct
Early Nov
Late Nov
Dec
Early Jan
Late Jan
Early Feb
Late Feb
March
Early Apr
Late Apr
Early May
Late May
Test Window

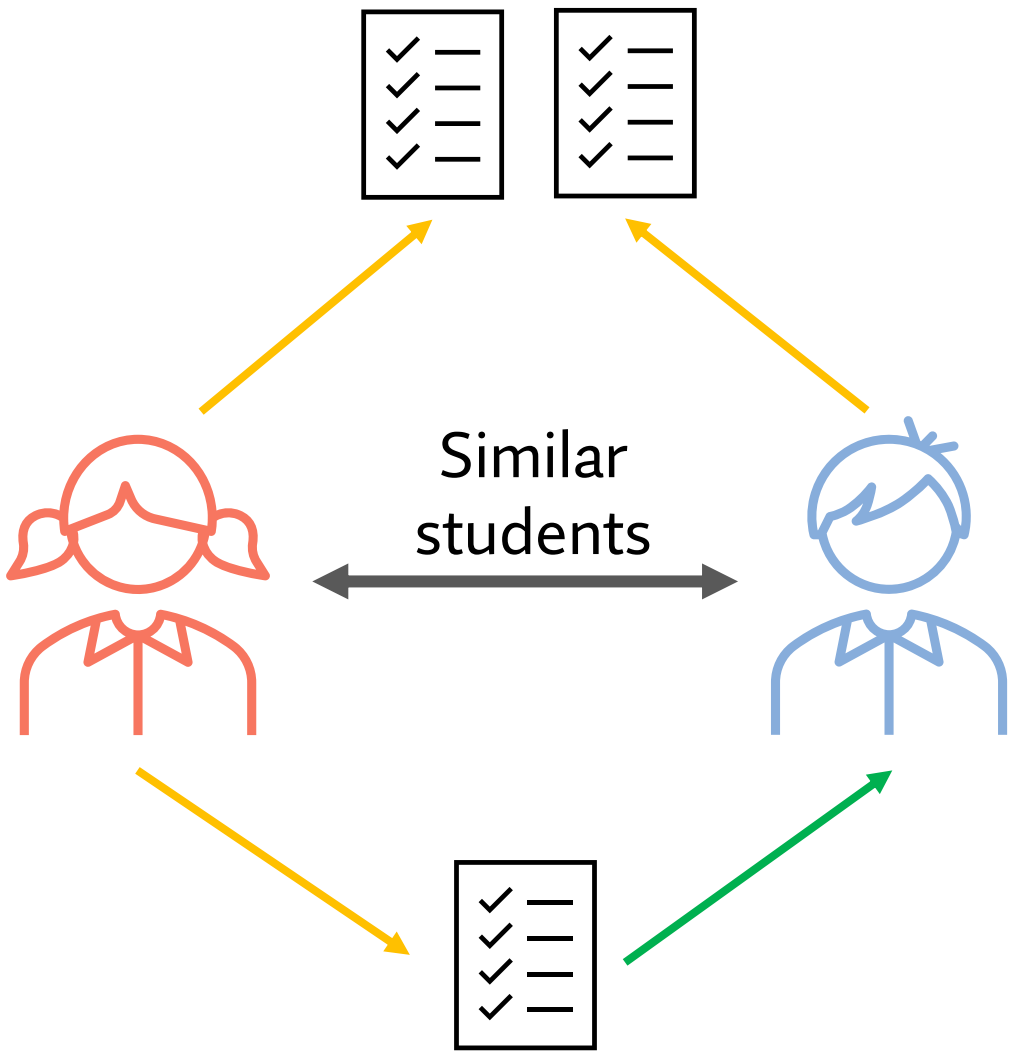

Reduce the number of tests
Maximize score change



Dijkstra's **Shortest Path First** Algorithm


Sep
Aug
Nov
May
Dec

Test taken by **both** students



Test taken by her, recommended to him

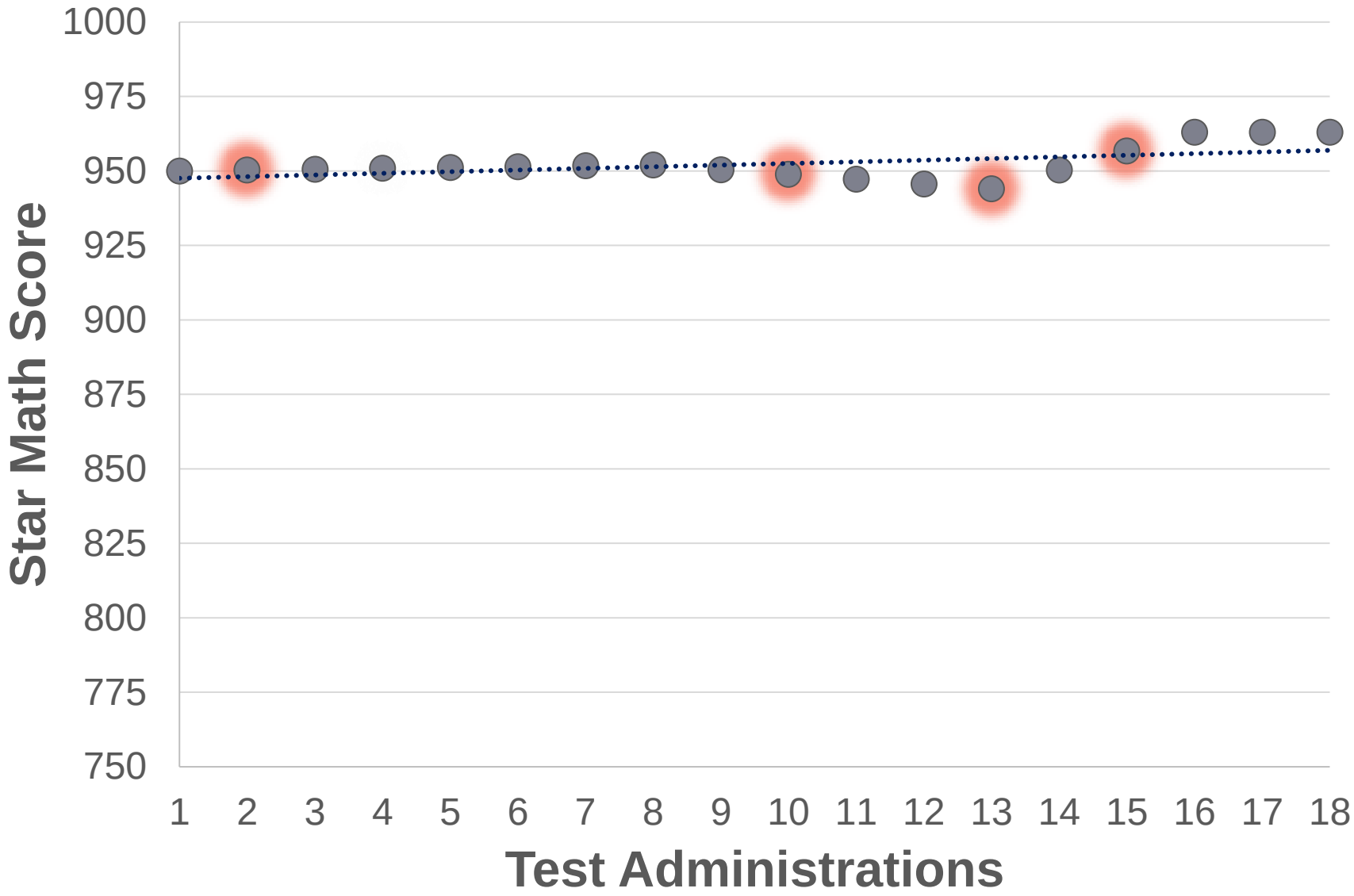
17 tests → 8 tests

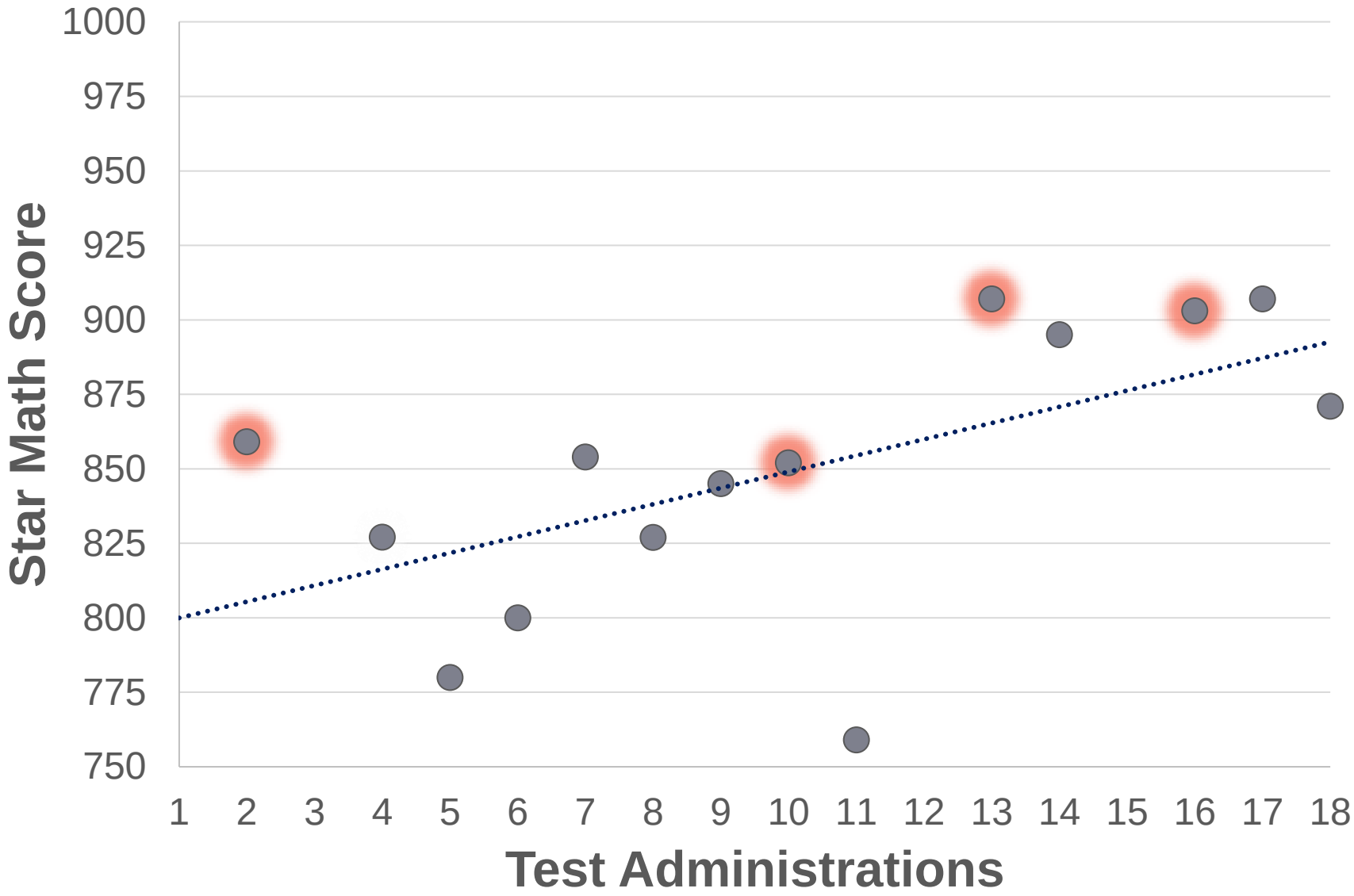
Grade 2

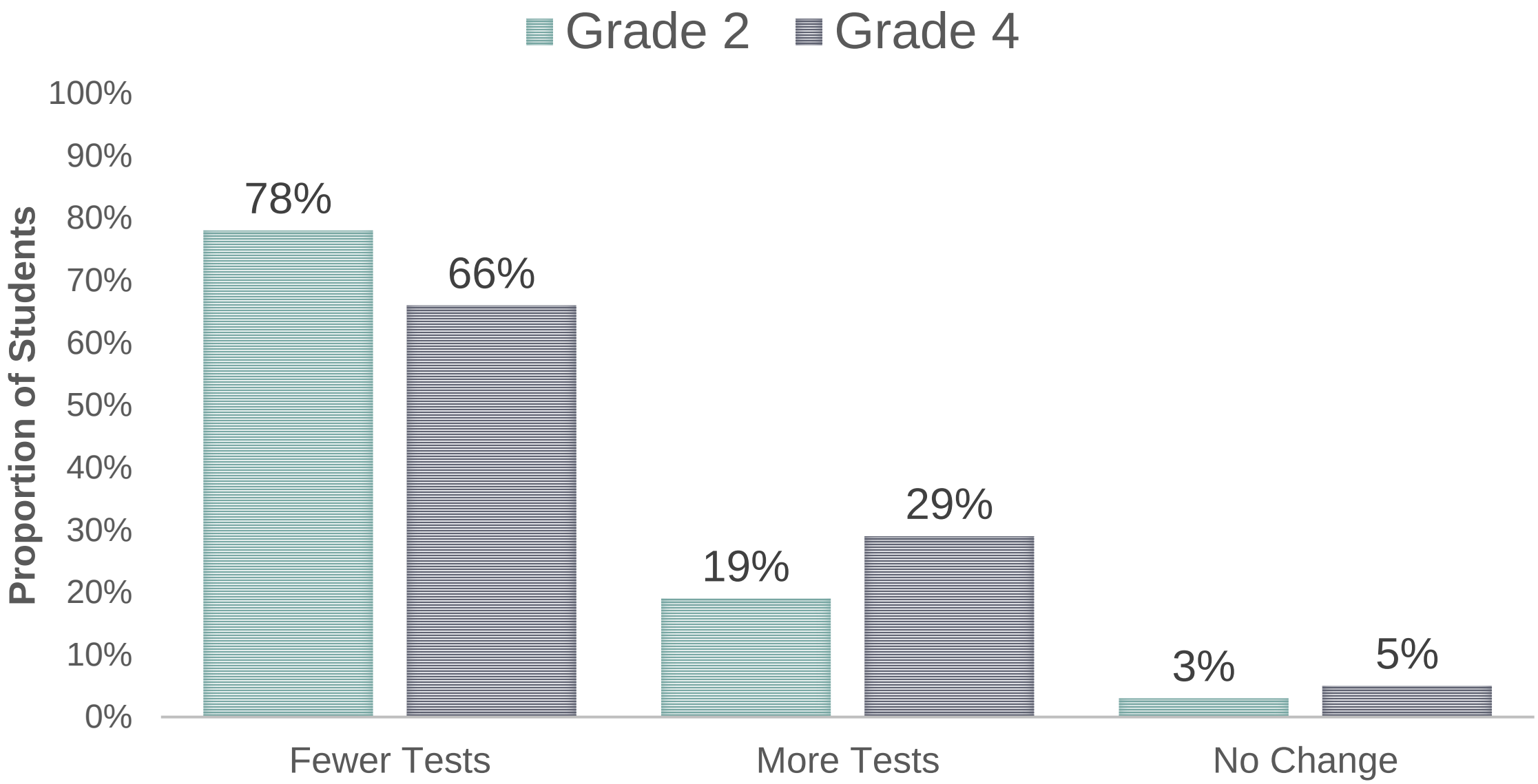
17 tests → 5 tests

Grade 4

Mission accomplished!


1000
975
950
925
900
875
850
825
800
775
750
Star Math Score
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Test Administrations


Star Math Score
1000
975
950
925
900
875
850
825
800
775
750
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Test Administrations


Grade 2
Grade 4
Proportion of Students
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
78%
66%
19%
29%
3%
5%
Fewer Tests
More Tests
No Change

frontiers in Education

ORIGINAL RESEARCH
published: 23 September 2020
doi: 10.3389/feduc.2020.572612

An Intelligent Recommender System for Personalized Test Administration Scheduling With Computerized Formative Assessments

Okan Bulut[1]*, *Damien C. Cormier*[1] and *Jinnie Shin*[2]

[1] Centre for Research in Applied Measurement and Evaluation, Faculty of Education, University of Alberta, Edmonton, AB, Canada, [2] Department of Educational Psychology, Faculty of Education, University of Alberta, Edmonton, AB, Canada

The introduction of computerized formative assessments in schools has enabled the monitoring of students' progress with more flexible test schedules. Currently, the timing and frequency of computerized formative assessments are determined based on districts and school authorities' agreements with testing organizations, the teachers' judgment of students' progress, and grade-level testing guidelines recommended by researchers. However, these practices often result in a rigid test scheduling that disregards the pace at which students acquire knowledge. Furthermore, students are likely to experience the loss of instructional time due to frequent testing. To administer computerized formative assessments efficiently, teachers should be provided systematic guidance on finding an optimal testing schedule based on each student's progress. In this study, we aim to demonstrate the utility of intelligent recommender systems (IRSs) for generating individualized test schedules for students. Using real data from a large sample of students in grade 2 (n = 355,078) and grade 4 (n = 390,336) who completed the Star Math assessment during the 2017–2018 school year, we developed an IRS and evaluated its performance in finding a balance between data quality and testing frequency. Results indicated that the IRS was able to recommend a fewer number of test administrations for both grade levels, compared with standard practice. Further, the IRS was able to maximize the score difference from one test administration to another by eliminating the test administrations in which students' scores did not change significantly. Implications for generating personalized schedules to monitor student progress and recommendations for future research are discussed.

Keywords: recommender system, formative assessment, personalized learning, progress monitoring, mathematics

OPEN ACCESS

Edited by:
Kasia Muldner,
Carleton University, Canada

Reviewed by:
Wei Shin Leong,
Ministry of Education (Singapore), Singapore
April Lynne Zenisky,
University of Massachusetts Amherst, United States

****Correspondence:***
Okan Bulut
bulut@ualberta.ca

Specialty section:
This article was submitted to Assessment, Testing and Applied Measurement, a section of the journal Frontiers in Education

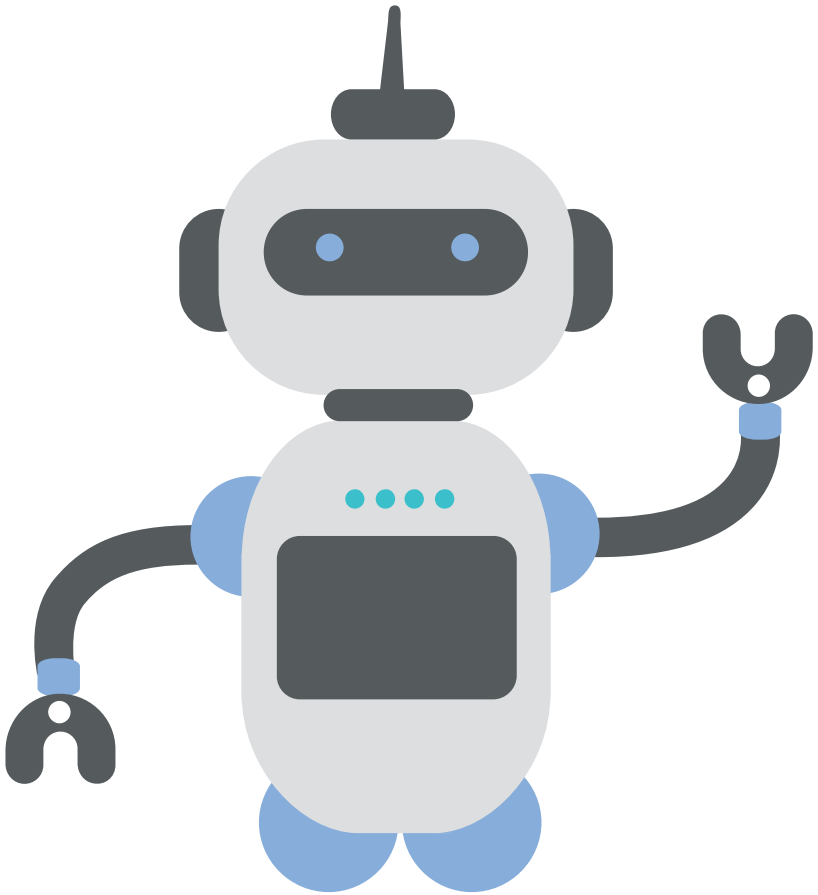
Received: *15 June 2020*
Accepted: *04 September 2020*
Published: *23 September 2020*

Citation:
Bulut O, Cormier DC and Shin J (2020) An Intelligent Recommender System for Personalized Test Administration Scheduling With Computerized Formative Assessments. Front. Educ. 5:572612. doi: 10.3389/feduc.2020.572612

INTRODUCTION

Classroom assessments allow K–12 teachers to evaluate student learning (i.e., monitor students' progress) and make a variety of important decisions about learning outcomes (e.g., producing feedback, assigning grades). Teachers use two types of assessments to evaluate student learning in the classroom: summative and formative (Black and Wiliam, 2009). Summative assessments are typically used *at the end* of an instructional period, such as a unit or a semester, to evaluate student learning relative to content standards or benchmarks. Quizzes, midterm exams, or a final project are

http://bit.ly/bulut2020



Thank You

bulut@ualberta.ca
www.okanbulut.com