Evaluating the effectiveness of adaptive testing in a power engineering technology course

Matthew Turner, Ph.D. and Rustin Webster, Ph.D.
School of Engineering Technology, Purdue University

Introduction

This study investigates the effectiveness of an adaptive testing strategy for automatically grading mathematically complex engineering problems using online platforms, specifically in a power engineering technology course. By comparing adaptive online exams with traditional handwritten exams, we aim to explore how fairly partial credit can be assigned in these settings.

Background and Purpose

Studies generally support the utility of online systems for math-based homework in aiding students' understanding of math-intensive subjects such as physics and statistics. Despite this, research specific to the fairness of online tools for exams has not been thoroughly explored. Traditional grading practices often emphasize not just the final answer, but also the process and understanding demonstrated by the student. However, many online systems are designed to evaluate only the final answer. This discrepancy highlights the need to explore how adaptive online exams can be used to fairly assign partial credit, particularly for numerical response problems.

Adaptive Testing

In this study, partial credit for online exams was managed by breaking problems into key-steps and using LON-CAPA to adapt test problems based on each student’s real-time performance (Figure 1). Initially, students attempt to solve the final answer directly with decreasing credit per attempt; if unsuccessful, the problem adapts to present key-steps sequentially.

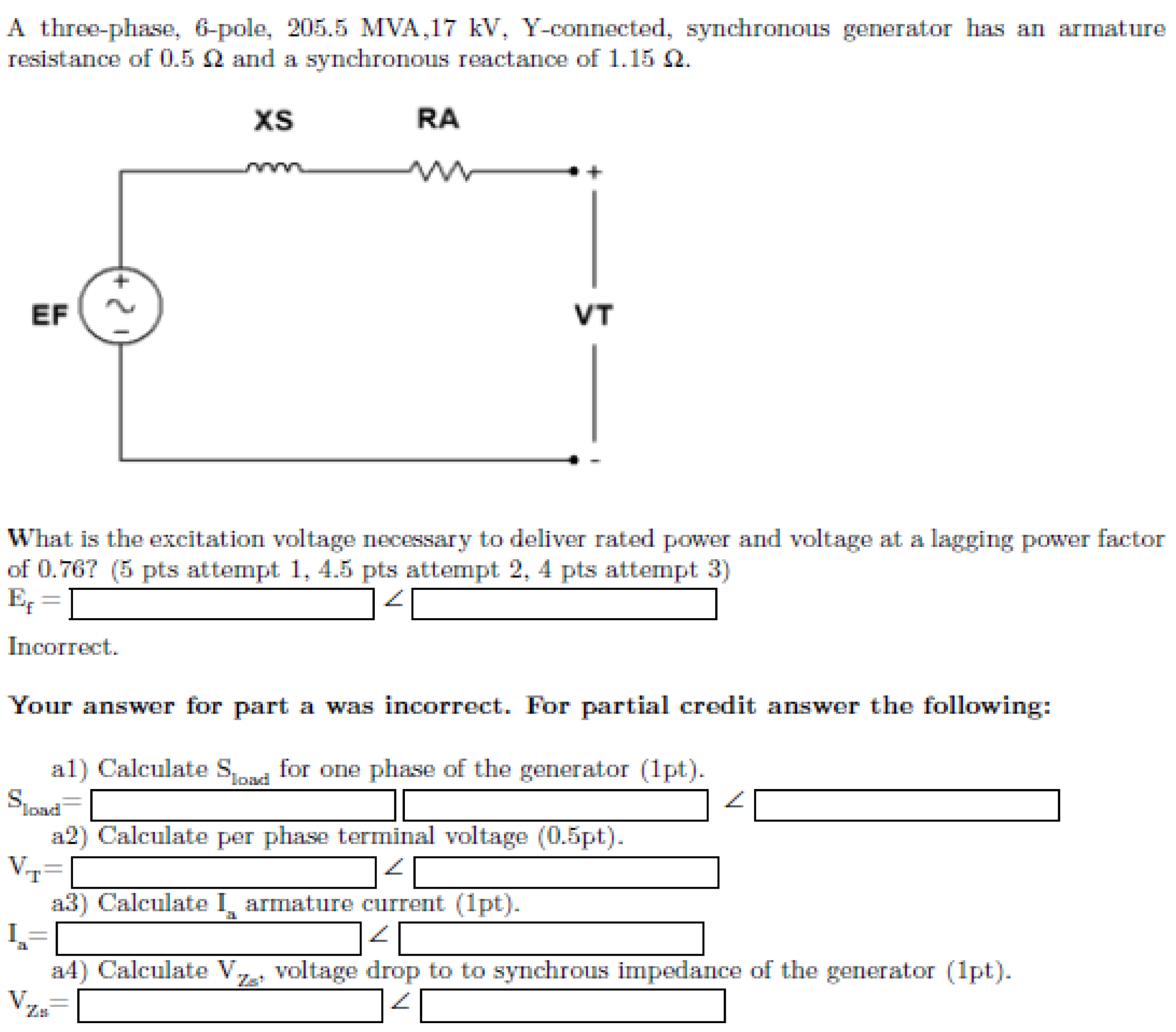

Figure 1. Adaptive LON-CAPA Assessment

Implementation

The study was conducted in a junior-level power engineering technology course over a four-year period. Assessments consisted of six in-class tests covering topics such as circuit analysis and electrical system models. In years 1 and 2, students (n=17) were assessed using traditional pen-and-paper testing. In years 3 and 4, students (n=12) were assessed using adaptive testing. All other academic instruction was held constant, including the use of computer-based homework using the online platform.

Results and Discussion

Preliminary findings are presented in the boxplot of Figure 2 for the average exam score by grading method. The plots exhibit similar distributions, as evidenced by their median lines being approximately equal and their interquartile ranges (IQRs) being comparable. This suggests that the central tendency and the spread of the data in both groups are closely aligned, supporting the conclusion that the approaches are comparably fair in assessing student performance. However, the whisker length for the traditionally assessed students is significantly longer than those for adaptive testing, indicating a greater range and suggesting more variability. This suggests a more punitive nature of computer-based testing for small mistakes significantly impacting scores. Additionally, the adaptive nature of the test may provide more guidance to students who are less familiar with the problem, leading to a longer lower whisker.

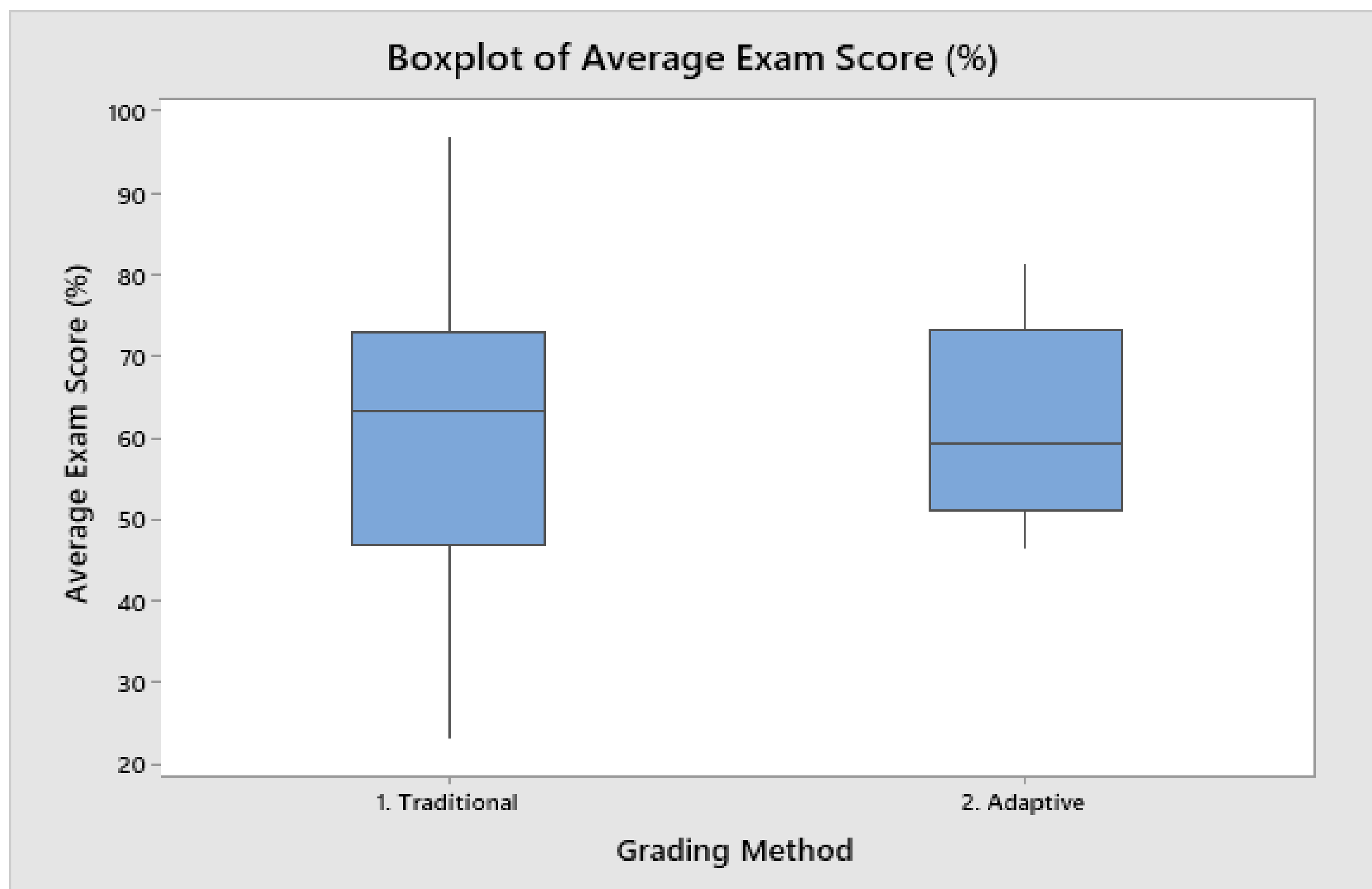

Figure 2. Average Exam Scores: Traditional vs. Adaptive

To address these issues, improvements should reduce penalties for minor errors, like missing a decimal point or incorrect calculations, to avoid unfairly impacting students for small mistakes. Additionally, the system should better identify when students truly lack understanding, rather than guiding them through problems, to ensure assessments accurately reflect their independent problem-solving skills, such as through more advanced branching logic.

Conclusion

This study compared adaptive online testing with traditional pen-and-paper exams in a junior-level power engineering course. Both methods showed similar average exam score distributions, indicating comparable fairness. However, traditional testing had greater variability, suggesting it may be more punitive for minor errors. Adaptive testing, with its step-by-step feedback, may provide excess guidance to students. To improve online assessments, focus on reducing penalties for minor errors and using advanced branching logic to better assess students' understanding and problem-solving abilities.