Examination of the Use of a CAT Learning Support System in Mathematics Department of High Schools

Yuki KANO
Tokyo University of Science, Japan
1721507@ed.tus.ac.jp

Hiroki OURA
Tokyo University of Science, Japan
houra@rs.tus.ac.jp

Yuki WATANABE
Tokyo University of Science, Japan
wat@rs.tus.ac.jp

Abstract

Item response theory (IRT) is a general statistical theory of how examinee performance relates to the abilities measured by test items (Hambleton & Jones, 1993). The computerized adaptive test (CAT) measures the ability of the examinees by determining the appropriate questions for individual examinees and efficiently changing the difficulty of subsequent questions (Yamashita, 2017). On the other hand, however, some students experience mathematics test anxiety (Putwain & Symes, 2010). Therefore, we developed a CAT system aimed at reducing mathematics test anxiety. In this study, we describe the outline of the CAT system we have developed and discuss how to use it in mathematics learning in high schools. Furthermore, we consider the feasibility of classes utilizing the analysis design development implementation evaluation (ADDIE) model and Gagné's nine teaching events based on the feedback content and ability values of the CAT in each session.

Keywords: Item Response Theory, Moodle, Computerized Adaptive Test, Mathematics Test Anxiety

Introduction

The Ministry of Education, Culture, Sports, Science, and Technology (MEXT; 2020) aims to realize an information and communication technology environment in which each student optimizes and fosters their qualities and abilities by providing one computer (e.g., laptops, tablets) per student as part of the global and innovation gateway for all (GIGA) school concept. One goal of the GIGA school concept is developing a school environment optimized for each individual student's learning. One of the promising methods being used to achieve the goal is the computerized adaptive test (CAT), which grasps the comprehension status of students. CAT combines computer-based testing and item response theory (IRT) to measure examinees' abilities by determining the appropriate questions for each examinee and efficiently changing the difficulty level of subsequent questions (Yamashita, 2017; Mitsunaga, 2017; Nakamura, 2011). The authors propose that IRT can help assess where each student is meeting and failing to meet their learning goals. The advantages of IRTs include evaluating multiple populations and taking different tests on the same scale. The ability values estimated using IRTs are reliable for measures such as the number of correct responses (Ueno, 2009).

Many students in high schools feel uneasy about mathematics tests. For example, Putwain and Symes (2010) reported more test anxiety in mathematics than in English or science. Fritts and Marszalek (2010) suggested that using CAT may increase test anxiety in students with little experience using computers. Mohd et al. (2019), however, reported that as compared to the traditional test, CAT can reduce test takers' anxiety about mathematics tests. Specifically, they

found that administering multiple CAT-based tests as practice before a mathematics test over three weeks or longer decreased mathematics test anxiety. Therefore, using CAT in daily classes at high schools could reduce students' anxiety about mathematics tests and make it easier for teachers to grasp each student's understanding of mathematics knowledge.

Therefore, in this study, we developed the CAT system to reduce students' mathematics test anxiety in mathematics at upper secondary schools. We discuss the effects of using this system in and out of class for learners and the feasibility of teachers utilizing the ability values and various scores given to test takers after using the system.

Overview of The Developed CAT System

Background of the system

This system targets the "development of educational methods" among the three priority goals of the "development of an individualized and optimized feedback system using mathematical data to promote the plan-do-check-act cycle of learning and educational development (MEXT, 2021)."

System Operating Environment

This system runs on Moodle, a learning management system to manage learners' progress and support class implementation (Fujiwara et al., 2007; Egi et al., 2017). It uses a two-parameter logistic model as the IRT model. The authors recommend taking the test on a PC or iPad.

System Overview

One feature of this system is that each question in Mathematics I, A, II, B, III, and C is designed and implemented by the New Courses of Study for Senior High Schools (MEXT, 2018). The questions are the same as those used in the Survey of Mathematics Basic Achievement for High School Students of Science and Mathematics developed by the Institute for Mathematics Education at Tokyo University of Science (2021). The answer format for each question is 5-choice (multiple choice). In the CAT system, each question is non-skippable, non-terminable, has a time limit, and posed one at a time. If an answer is interrupted owing to a technical error or communication error, the user can restart the interrupted question. Furthermore, if a student takes the same test more than once, the test will start from the student's ability value at the time of the previous test.

The test flow comprises an initial set of questions with varying degrees of difficulty at the beginning of the test and a provisional estimate of the examinees' ability value. Suppose the item parameters of the questions are known. In that case, the minimum and the maximum number of questions, including the initial question and the percentage of questions in each field, should be set. The next question should be from the field that deviates the most from the current percentage of questions in each field. On the other hand, if the item parameters of the questions are unknown, a question estimating the item parameters is presented for every random number of questions estimating the ability value. The end of the test is determined when the standard error of the ability value falls below the threshold set for the item pool, when the examinee has answered the minimum number of questions required for ability estimation, when the time limit has elapsed, or when the end date and time of the available test period have passed.

When the CAT system answers all questions correctly or incorrectly, the CAT system estimates the ability value by Bayesian estimation based on the answer data up to that point. Once even one correct or incorrect question is answered, the ability value is estimated using maximum likelihood estimation. The test taker then receives a score that includes the test result (on a 5-point scale of estimated ability), the standard normal distribution of the item pool,

the percentage of correct responses (%) for the test as a whole and for each domain (% and radar chart), and the number of attempts. In addition, the CAT system uses the last examination data for each examinee as the aggregate data. For authors and teachers, group averages are checked in addition to those of individual examinees. Additionally, they can also check other test units.

How to use the CAT system

In this section, we consider how to use the CAT system. For example, if a high school student owns a computer, they can use the system inside and outside class. In class, the system can give quizzes at the beginning of each class. Outside of class, the system can give an out-of-class assignment at least once a week. Merely by using the system over a long period, it is possible to verify changes in ability values and mathematics test anxiety. Furthermore, we believe that it is possible to design classes that use information obtained through this system, such as ability values. Specifically, we believe that teachers can use the information for individualized instruction and grouping students in classes.

Future Work

This study has shown the overview of the developed CAT system and how it is used at school. In the future, an evaluation study is going to be conducted to examine the CAT system in the teaching of mathematics in high school. Specifically, we are going to implement this system at a high school where teachers and students have their own iPads. Specifically, we will first create and show students a video or slide presentation of how to use the CAT system. Next, students will use the system outside of class every week in two groups; one group uses the CAT system, and the other does not. We will switch the groups every month and evaluate the changes in the students' sense of anxiety and ability in mathematics tests by comparing the results of pre-, mid-, and post-tests. Questionnaire surveys will also be administered at pre-, mid-, and post-stages to measure students' mathematics test anxiety. In addition, students will be asked to write about their impression of the CAT system in the post. Lastly, we will also design, practice, and evaluate lessons utilizing information obtained from this system using various instructional design theories (the ADDIE model and Gagné's nine teaching events).

References

Egi, H., Hayashi, K., & Tsujisawa, T. (2017). A survey of teachers' awareness of systematic utilization of learning management systems. *Transactions of the Japan Society for Educational Technology,* 40, Suppl., 189-192. (in Japanese)

Fritts, B. E., & Marszalek, J. M. (2010). Computerized adaptive testing, anxiety levels, and gender differences. *Social Psychology of Education,* 13, 3, 441-458.

Hambleton, R. K., & Jones, R. W. (1993). Comparing classical test and item response theories and their applications to test development. *Educational measurement: issues and practice,* 12(3), 38-47.

Fujiwara, Y., Ohnishi, H., & Kato, H. (2007). Trends and issues in research on mutual assessment among learners. *Journal of Media Education,* 4(1), 77-85. (in Japanese)

Institute for Mathematics Education, Tokyo University of Science (2021). Mathematics Proficiency of High School Students NOW XVI -2020 Report on the Survey of Basic Academic Proficiency-. *Kagaku Shinshinsha*

Mohd Ali, S., Norfarah, N., & Ilya Syazwani, J. I. (2019). The effect of Computerized-Adaptive Test on reducing anxiety. Towards math tests for polytechnic students. *Journal of Technical Education and Training,* 11(4)

MEXT (2020). The GIGA School Initiative. Retrieved July 3, 2022, from https://www.mext.go.jp/kaigisiryo/content/20200706-mxt_syoto01-000008468-22.pdf (in Japanese)

MEXT (2021). "Plan for the Advancement of Education at Universities and Colleges of Technology Utilizing Digital Technology" Outline of Efforts during the Implementation Period (2). Retrieved July 3, 2022, from https://www.mext.go.jp/content/20210630-mxt_senmon01-000016115_2.pdf (in Japanese)

Mitsunaga, Y. (2017). *What do Tests Measure: The Concept of Item Response Theory* (pp.186). Kyoto, Japan: Nakanishiya Shuppan. (in Japanese)

Nakamura, Y. (2011). The possibility of computer-adaptive testing. *Japanese Language Education,* 148, 93-110. (in Japanese)

Putwain, D. W., Connors, L., & Symes, W. (2010). Do Cognitive Distortions Mediate the Test Anxiety-Examination. *Performance Relationship? Educational Psychology,* 30(1), 11-26.

Ueno, M. (2009). e-Testing: Advanced theory and technology. *Journal of Educational Systems and Information Science,* 26(2), 204-217. (in Japanese)

Yamashita, Takuya (2017). Prototyping Computer-Based-Testing (CBT) questions. *Chemistry and Education,* 65(7), 334-337. (in Japanese)