Educational Measurement and Educational Evaluation of Decision Making based on IRT

Fei Wang
Department of Adult Education, Wen Shan University, Wen Shan City, 663000 China

Keywords: IRI; Teaching Evaluation; Evaluation; System Design

Abstract. Teaching evaluation is an important part of the teaching process, which is also the most important means in the quality assurance system of education. As an important tool for teaching measurement and evaluation, the test’s five functions, namely, unique assessment, testing, diagnosis, feedback and incentive, are directly related to the accuracy and objectivity of teaching evaluation, but also affect the good or bad teaching quality.

Introduction
In the traditional teaching, we often emphasize the excessive assignments tactics, many of which are the repetitive labor and stay on the questions with the same difficulty, which will only increase the amount of the students’ workload, but will not improve the efficiency of learning. In the distance education, because there is no timely guidance from the teacher, students can only choose the questions to practice by their own subjective judgment, while in the adaptive learning, exercise and testing are an indispensable component, where a support system is required to provide learners with an assessment for learning performance and to provide appropriate learning recommendations under the condition when there is free of a teacher's participation, but for the adaptive exercise, no perfect model has been put forward. The current exercise support system has two kinds: the first is the test question beyond the textbook, and the second is the exercise software with the computer as the hardware basis. The former can not evaluate the performance of the learners, or provide the appropriate questions (called the project in the IRT theory), while the method for extracting questions in the latter is usually a transfer of the questions on the textbook, in which the algorithm for the extracting questions is too simple, which is usually extracted randomly after a certain range is specified.

Basic Principles of IRT

As we all know, the purpose of a test is to estimate the internal latent traits which can not be measured directly for the testee with the test scores, so as to deduce and evaluate his or her development level. IRT is based on the theory of latent traits, which describes this relationship by building a certain project response models. The most important feature of IRT is that it has found an item characteristic curve and describes and approaches it in a variety of mathematical expressions (namely, mathematical models). Obviously, different mathematical models have different degrees of approximation to curves, and also contain different numbers of parameters. The mathematical model of the item characteristic curve usually contains the parameters on two aspects:

1. Item parameters that describe the characterization of test items;
2. A latent trait also called capacity parameter that describes the characteristics of a candidate.

Based on these parameters, the quality of the test item can be evaluated, and the response of the individual candidate to any test item can also be measured, so as to estimate the number of latent traits, namely, the ability or trait that is owned by a candidate when the test item is completed. Theoretically, it has effectively solved the problem where the classical test theory can not establish the function relationship between the candidates’ scores and the test item parameters.
Evaluation System Requirements based on IRT Decision

The adaptive network test system is an important part of the whole teaching process, the purpose of which is to serve for the teaching, so in the design of the system, we must fully consider the hardware and software environment of the system and the teaching function requirements. Based on the analysis of the self characteristics of adult education, the theory of test and the development of network technology before, we believe that the adaptive test system of the scientific, safe and applicable network shall meet the following requirements:

1. Allow the users (students or teachers) to use the browser to access the system through the Internet to achieve various functions;
2. Facilitate the maintenance and management of the database;
3. Have a perfect test control function. The system shall have a scientific test generation strategy, a complete test process control and automatic evaluation function of test questions;
4. Perfect data record, statistics and analysis functions;
5. Have a good user interface to minimize the factors that affect the results of the test;
6. Have a good scalability;
7. Have good security and confidentiality mechanism.

Evaluation System Design based on IRT Decision

The question bank is an important technical means to ensure that the examination or test has a higher quality and stable level and can better achieve the prediction purpose. It is not a simple collection of some questions, but a scientific combination of some questions that have been specially treated and have larger information amount. The quality of a question bank depends on the scientific rigor of its organization, the breadth and typicality of its content, the predictability of the candidate's ability and the practical feasibility, etc. For the CAT test, the question bank is more a basis for effectiveness and efficiency of the test.

The question bank used by the system is based on C Language Programming edited by Tan Haoqiang and published in Tsinghua University Press. In order to cooperate with the system algorithm, the question type only uses the multiple-choice question, but has reserved the interface for the future use of multiple courses and various question types in the design.

The question bank is a limited set suitable for a certain test target and with relevant necessary parameters and a large number of excellent test questions, which is the memory bank of a large number of excellent test questions and the source of the main data of the test system. The main function of the question bank’s management module is to complete the establishment and maintenance of the IRT question bank, which mainly includes:

1. Input of test questions. The questions input shall be predicted or evaluated by the expert before confirming the IRT parameters of questions with statistical method (including discrimination, difficulty and guessing parameters, etc.);
2. Maintenance of question bank. Include basic management functions such as query, addition, deletion and update of questions, etc.

The main function of the system management module is to carry out the daily management and the security of the system, including:

1. Registration management: management of the user's application, account password management and user cancellation. Only the students in the database have the authority to use the system to test and query the relevant information; for the teachers who are granted authority, the authority to enter the examination of questions is divided into three levels:
   Level 1 has all authorities like browsing question bank, modifying and entering the questions, browsing the test result analysis.
   Level 2 has the authority to browse the question bank and test result analysis, but has no authority to modify and input.
   Level 3 can only check the test result analysis of the taught class, but has no authority to browse and modify the question bank.
The user’s authority is assigned by the administrator in unification. The authority of the course teacher and student’s data of taught class is input by the administrator in unification and other students can apply for registration.

(2) Data management: distribute and manage the authority of users at all levels, who use the system, and set the system operation number (such as the test time and amount of questions etc.). Password, test result and answer, etc. are transmitted on-line, therefore, it is necessary to carry out encryption processing, but the concrete encryption method will not be discussed in the Paper.

The main function of the test management module is to complete the online test and the analysis evaluation and query of the test result analysis.

(1) Quizzes and tests. Include the adaptive test based on item reflection theory, which is by chapter, and the adaptive test of the whole course, and provide the traditional simulation test based on classical test theory. The question-selection strategy algorithm of adaptive test is based on the three-parameter logistic model of IRT; the paper generation of the traditional practice test form is based on the classical test theory and in accordance with the knowledge points, difficulty, distinction and other conditions.

(2) The system in the test automatically records the students ‘corresponding test such as the use of the questions, the knowledge points of the test questions, the candidates' answers and the candidates 'test results of this time, etc., which are used as the basic data for the analysis evaluation of the students themselves and the teachers' teaching result analysis, and as one of the basis for question selection of adaptive test.

(3) Statistical Analysis. The system analyzes the user’s test to provide teaching analysis and organization of teaching activities for the students and teachers, which is mainly to make a statistics and analysis on each test and the answers composing the test and an analysis on student’s learning, learning ability and learning result according to historical records, including:

According to the students' performance in the test process, the students' learning ability is analyzed. That is to say, carry out interaction between the question bank system and students through the test interface, in which the computer records and analyzes student’s various test response information. After analysis, a lot of valuable information about the teaching process is obtained. Because each question in the test examines one or more of the knowledge points and the cognition degree, it is possible to understand the scope and degree that are obtained by a student on some knowledge point by analyzing the performance of the student on the questions and to get the mastering of knowledge point and capacity development. The analysis results can guide the student's individual learning behavior.

Through the horizontal statistics comparison, carry out the statistical analysis on test of all students and on answers to questions used to get the information of teaching result on each knowledge point in course and information of individual student’s mastering and ability on course and display by means of graphics.

Carry out statistics on the use of questions and warn the abnormal questions, such as the test with distinction as 0. Meanwhile, provide all candidates’ highest score and lowest score, the number of people in each score section, average score, standard deviation and other statistical calculations for the traditional practice test, which shall be shown by graphics to intuitively show the distribution of student performance, so that teachers can intuitively and vividly as well as accurately understand the test result.

References


