PERSONALISED LEARNING THROUGH COMPUTER ADAPTIVE TESTING

G. USHA, Ph.D. Scholar, Madurai Kamaraj University, Madurai
Dr. R. VIJAYA, Guide, Asst. Professor & Director i/c, UGC - Human Resource Development Centre, Madurai Kamaraj University, Madurai.
Dr. M. RAJESHKUMAR, Co-guide, Asst. Professor, Centre for Educational Research Madurai Kamaraj University, Madurai.

Abstract

The traditional design of a test is fabricated into fixed forms. This paradigm has a number of problems; some crystal clear and some more indefinite. Also the most disadvantage of the traditional one is that it is quite inefficient. Students with low proficiency are exhausted and discouraged by attending most of the items that are very difficult. Students with high proficiency are not appropriately challenged because most of the items attended are very easy. With the development of computer, Computer Based Testing (CBT) has been much sought-after current testing system. There are two types of CBT namely linear and adaptive. Compared to linear tests, Computer Adaptive Testing (CAT) technology requires fewer test items to arrive at a more accurate estimate of students. This paper highlights personalized learning of students through CAT, the importance of adaptive testing and the potential of CAT for assessing students' performance and progression. Also, it focuses on the present CAT platforms available for more precision and efficiency.

Keywords: *Personalized learning, Computer Adaptive Testing, ability level*

Introduction

Technology empowers teachers to measure the results of their lessons in a more effective way. It is a challenging task for teachers to assess new teaching methods and to find how different content and other variables impact on the learning experience and the

results obtained by students. Also the results help teachers analyse the course in between when the scores or results are below the expected level. Computer-based technology is renovating the whole teaching-learning process. In particular, they influence the creation of types of tests and their assessment in learning.

Personalized Learning

Personalized Learning focuses on the learner. The traditional approach emphasizes "one system fits all". This is being replaced with a new approach of "selecting what fits best for the student". This assists to meet the individual needs and interests. A vital part is being played by the technology and digital skills. The arrival of technology has brought about changes in the way of teaching and learning. Also, the communication between the students and teachers has been made simple. Online course content, classes and lectures, software has been brought into classrooms by technology for students with special needs. There are so many benefits for teachers and students. The parents are also welcoming new technologies.

Need of a New Form of Assessment

When a standardized test is conducted for a student with high proficiency level, there is no great difference in attending it. Usually the student gets all the questions right and would end up with great commendation. On one side, the students with low proficiency level get low scores when compared to the others. But on the other side, the plight of the high proficiency level students is very pathetic because they are not treated with rich learning experiences. The status quo of them is that they don't deserve the same items to be administered for them. The test scores would have given a picture that these students are far off with the low proficiency level students but the real truth is that they are not made clear about what they don't know and what type of instruction they required. Since the students are of diverse nature, a new form of assessment is now available to engage and assess the students according to their ability.

Students have to demonstrate their abilities to use what have been learnt because the affection grows in performance-based testing. For this, new ways of performance assessment are to be developed. The computer based technology smoothen the path of

test administration, scoring, data analysis and reporting. Computer Adaptive Testing technology needs fewer test items to reach at a more accurate estimate of students unlike conventional tests. No two children learn the same. Computerized adaptive assessments are thus developed to test differently and allow teachers to see their students as they are each with their own ground of knowledge. With flexible delivery options, these assessments can scale to fit our needs. The outcome will be fruitful for the students, teachers and administrators.

Adaptive learning

Its focus is on designing a learning system which is based on the individual needs of the students. It is student-centred. The technology empowers to make changes very easily in course content or assessment. There is software available to keep trace of student progression and make adjustments according to those observations.

Computer Adaptive Testing

Computer-based technology is renovating not only the structure and content of the curriculum of schools but also the whole teaching-learning process. In particular, they influence the creation of types of tests and their assessment in learning. "With the furtherance in information technology and psychometrics, paper and pencil tests can be changed into computerized adaptive testing format." **(Lord, 1980; Weiss, 1976; Wainer et al., 1990)**. It must be signified that "computer adaptive testing is unlike computerized administrated testing which usually makes a mention of a system that casually selects a test item or a subtest from a pool of items without having consideration on the ability of the students" **(Beevers et al., 1995)**. The supremacy of CAT over conventional test is that it enables more potent and precise estimation of traits of the students.

With CAT, students take tests that best suit their ability levels. This happens just by adjusting the stimuli to the student based on his/her preceding response. The intensity of difficulty of the subsequent stimulus is selected so that it is not too easy or too difficult for the examinee. The first item is a medium-difficulty one because the abilities of the students are unknown. The second stimulus is adapted to the ability level of each examinee. The computer thus adapts to the student's ability level and evaluates his/her

knowledge. Jettmar and Nass cite CATs as a distinct case of intelligent user interfaces and also added that the performance of the students is quietly monitored and the difficulty level of the questions are adapted accordingly. Brusilovstry describes CATs as an element of a paradigm shift within educational software development from “one size fits all’ to one able to provide maximum levels of interaction and personalization.

At the end of the test, no one is likely to get all answers wrong and scores zero mark; the less competent students would find some items that they could solve and hence retain their interest and motivation in the subject. Neither anyone is likely to get all answers correct and scores full mark; thus even the top students understand that there are rooms for improvement. The number of correct items for two examinees may be same but their resulting score relies upon the difficulty level, discrimination and pseudo guessing. **(Hambleton & Swaminathan, 1985)**.

All students have the capacity for learning. They begin their learning at different levels of comprehension, learn in different ways and show different levels of enthusiasm, interest and motivation to a given task or topic. The test results are more accurate and meaningful now through computerized adaptive testing which is a new assessment paradigm shift.

Models of Adaptive Test

Linn and Rock (1968) and Lord (1971) introduced two-stage testing, which was later developed by Earl, and it is used along with item response theory to develop two different sets of placement tests for textbook series by **James Dean Brown (1981).** Two-stage testing are very helpful in norm-referenced testing situations (typically for proficiency or placement purposes) for saving the time as well as to avoid answering many items which are very easy or too challenging for them. The two-stage testing can be applied to any test in which the students start by undertaking a short routing test with a number of difficulty levels. Here the answer scores help to select the path of the tests they should take (see Figure 1.1). Their final result is based on standardized scores that are matched across the High Difficulty, Medium Difficulty and the Low Difficulty measurement tests.

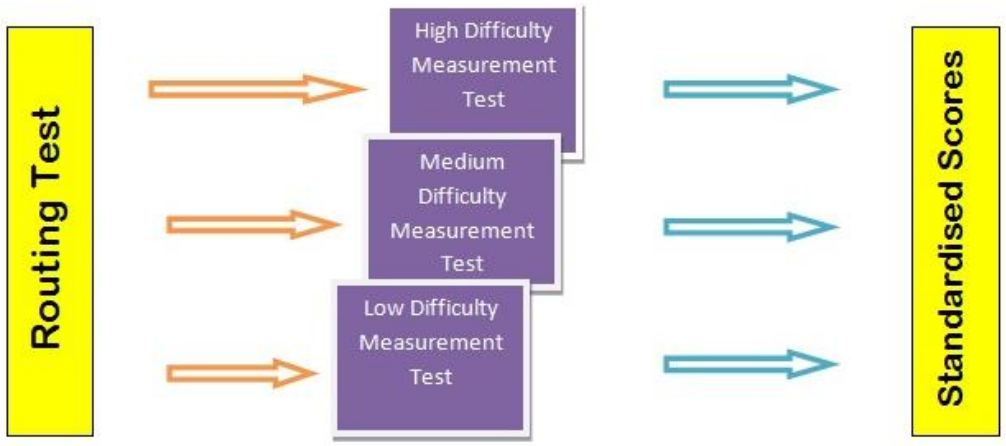

Figure 1.1: A Model of Two-Stage Testing

Computer adaptive testing shows difference on the concepts which were first developed for two-stage testing. In CAT, the students are guided by a routing test in which the test items exactly match at their ability levels. With the help of the information from the routing test the computer chooses items for each student's ability level. Every student takes a test which is different. The test is shorter and more precise than a two-stage measurement test. Moreover, CAT requires (a) a large itemised bank to be field tested and analyzed, (b) the test developer should have rich experience in item-response theory statistics and (c) the test developers have significant computer programming knowledge unlike in two-stage testing.

A multistage adaptive test (MST) has multiple modules which are administered at different stages during the administration of test based on students' prior answer or response to the test item. A module is a set consisting of items that are administered as a test unit. A traditional computerized adaptive test adaptively selects individual items whereas a MST adaptively selects individual modules at each stage.

Computer-adaptive sequential testing (CAST) was built as an integrated framework for high-stakes multi-stage computer-adaptive and mastery tests **(Luecht & Nungester, 1998)**. At present, many applications have adopted CAT designs because it gives the more precise measurement for all examinees than the traditional one. Another

important CBT model is MST. There is a difference between MSTs and CATs. The examinees always receive a set of pre-assembled items that are matched to their provisional ability estimates in a MST **(Hendrickson, 2007)**, whereas in a CAT, a single item alone is selected to suit the estimation of their ability..

A MST has several stages. There are several bundles of items pre-assembled before test administration within each stage. Each module is created based on item difficulty levels. This gives information for a particular ability level that can be maximized. Once all of the modules are pre-assembled, the modules in different difficulty positions can be bundled together in a unit which is called “panel” **(e.g., Luecht, Brumfield & Breithaupt, 2006; Zenisky, 2004)**. A MST design includes three stages and ten parallel panels. Each panel includes seven modules of varying difficulty levels. In addition to this, there are seven available pathways for each examinee; they are 1M+2H+3H, 1M+2H+3M, 1M+2M+3H, 1M+2M+3M, 1M+2M+3E, 1M+2E+3M and 1M+2E+3E. Here, “1”, “2” and “3”represent first, second and third stage respectively. The E, M, and H represent easy level, medium level and hard level respectively.

Various different statistical models are used to represent both item and test taker characteristics. There is no assumption of each item is equally difficult. This distinguishes IRT in which "All items are assumed to be replications of each other or we can say items are considered to be parallel instruments". The item response theory treats the difficulty of each test item as information to be incorporated in scaling items.

The difference in these estimation approaches in the estimation is based on the number of parameters. The most wide-spread approach Rasch-model, (the 1-parameter-logistics-model) (or), only requires to determine the difficulty level of each item. The remaining parameters–discrimination and guessing remain fixed here. 2- and 3-parameter-logistics are give a more exact and faster estimation of an examinee’s proficiency level, but require to find the extra parameters for each test item. Under IRT, a lot of research work has been done and using them as base, the Researcher has built up the present study on Computer Adaptive Test with three

Conclusion

Computerized adaptive testing (CAT) is an innovative method of student testing. The difficulty level and the number of items are tailored to the students' ability level. Students who are smart, risk more difficult items. The interaction of the test and intelligent algorithm makes CAT more beneficial. Software and expertise are the two important dimensions to be taken care of helps in the successful implementation of CAT.

References

- Beevers, C.E., McGuire, G.R., Stirling, G. & wild, D.G. (1995). Mathematical Ability Assessed by Computer Computers *& Education, 25(3),* 123-132.
- Hambleton, R.& Swaminathan, H. (1985). Item Response Theory: Principles and Applications. Boston, MA: Kluwer Nijhoff.
- Hendrickson, A. 2007. An NCME instructional module on multistage testing. Educational Measurement: Issues and Practice, 26, 44–52.
- James Dean Brown (1981). What is two-stage setting? *Shiken : JALT Testing and Evaluation SIG Newsletter Vol. 5 No. 2. June 2001 (p. 13-16),* University of Hawaii at Manoa.
- Lord, F. M. (1971). A theoretical study of two-stage testing. *Psychometrika, 36,* 227–242.
- Luecht, R. M. & Nungester, R. J. (1998). Some practical examples of computer adaptive sequential testing. *Journal of Educational Measurement, 35, 229-247.*
- Luecht, R. M., Brumfield, T. & Breithaupt, K. (2006). A testlet-assembly design for adaptive multistage tests. *Applied Measurement in Education, 19,* 189–202.
- Meyer & Zhu,(2013). Fair and Equitable Measurement of Student Learning in MOOCS: An Introduction to Item Response Theory, Scale Linking, and Score Equating : *Research & Practice in Assessment*, vol.8 pp. 26-39.
- Scalise, K., & Allen, D. D. (2015). Use of open-source software for adaptive measurement: Concerto as an R-based computer adaptive development and

delivery platform [Review of the test Concerto]. *British Journal of Mathematical and Statistical Psychology*, *68*(3), 478–496.

- Triantafillou, Georgiadou, & Economides, 2008. The design and evaluation of a computerized adaptive test on mobile devices. *Computers & Education*, vol. 50, no. 4, pp 1319-1330.
- Wainer, H. et al. (Ed., 1990), *Computerized Adaptive Testing: A Primer*. Hillsdale, New Jersey: Lawrence Erlabum Associates, Publishers.
- Weiss, D.J. (1976). Adaptive testing research in Minnesota: Overview, recent results, and future directions. In C.L. Clark (Ed.), *Proceedings of the first conference on computerized adaptive testing, 24-35*. Washington, DC: United States Civil Service Commission.
- Zenisky, A. L. (2004). *Evaluating the effects of several multi-stage testing design variables on selected psychometric outcomes for certification and licensure assessment*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.