Olympiads
in Informatics

11

OLYMPIADS IN INFORMATICS

Volume 11   2017

Selected papers of
the International Conference joint with
the XXIX International Olympiad in Informatics
Tehran, Iran, 28 July – 4 August, 2017

OLYMPIADS IN INFORMATICS

The journal Olympiads in Informatics is an international open access journal devoted to publishing original research of the highest quality in all aspects of learning and teaching informatics through olympiads and other competitions.

http://ioinformatics.org/oi_index.shtml

Foreword

The International Olympiad in Informatics (IOI) community gathers each year during the olympiad and has the opportunity to communicate during this worldwide event. Each country has many things to present and discuss. The national olympiads do not exist in isolation, and the papers from the 11th IOI conference show how similar problems arise in different countries and different environments.

The IOI Journal is focused on the research and practice of professionals who are working in the field of teaching informatics to talented students. The format for the Journal follows three tracks (since the 5th volume): the primary section of the journal focuses on research, the second section is devoted to sharing national experiences, and the last section presents book reviews and other related information.

In this volume, we publish a detailed research on the development of the curricula for Computer Science education in schools, in the paper "On a Methodology for Creating School Curricula in Computing" written by professors Krassimir Manev and Neli Maneva. The authors propose an idea for methodology for the development of school curricula in computing. These ideas were raised during the IOI Workshop, held in Bitola, Macedonia, in April 2015. The main goal of the Workshop was to review the current state of teaching computing in schools in the presented countries and to propose some ideas for development of an International School Curriculum in Computing, based on the experience of the IOI community.

The paper "Learning and Teaching Algorithms Design and Optimisation Using Contests Tasks" by Sébastien Combéfis and his colleagues presents training materials built from contest tasks to teach and learn how to design algorithms that solve concrete and contextualised problems. The first learning modules will be built thanks to a pedagogical device that will be deployed during the 2017–2018 academic year. The next step of this work, after the creation of the first modules, will be their evaluation in real conditions, that is, using them to teach algorithm design and optimisation related aspects and topics to interested learners.

The IOI conferences sustain the aim to become a part of, and to bring in, the wider pedagogical community. Each year there are typically one or two papers devoted to this purpose. In this volume we have a detailed paper titled "Learning Trajectory of Item Response Theory Course Using Multiple Softwares" written by Henri Retnawati from Indonesia. The paper describes learning trajectory of item response theory (IRT) that has been integrated to use multiple software and obstacles in an advanced IRT course. This design research study is of validation-study type with qualitative approach. The learning trajectory that had been hypothesized was based on the researcher's experi-

ence in lecturing and in identifying the necessary requirements for each chapter in the course.

Michael Dolinsky presents a paper "A New Generation Distance Learning System for Programming and Olympiads in Informatics". He proposes a new system, which supports teaching and learning of an arbitrary programming language, provides personal approach, and guarantees quality of knowledge and skills of graduates.

Antti Laaksonen discusses competitive programming in teaching university algorithm courses, and Dennis Komm presents an introduction to runtime analysis for an SOI workshop.

Some of the other papers in this volume deal with improving teaching and learning computer programming at secondary level, teaching graphs for students at lower secondary level, constructing problems in the sense of the IOI in Discrete Mathematics and Theoretical Informatics, and training world-class professionals in ICT.

We also understand the need for continuing to share our national experiences – our problems are common problems. In the second part of the volume, Bulgaria, Japan and Kazakstan present thoughtful analyses on the situation in their countries concerning informatics olympiads. IOI community members Bakhyt Matkarimov, Gregg Lee, Margot Phillips and Eljakim Schrijvers review IOI host guidelines and general aspects, which we hope will be of interest to our readership.

Finally, Antti Laaksonen presents a new book on Competitive Programming, and Krassimir Manev and Biserka Yovcheva announce the first European Junior Olympiad in Informatics.

As always, thanks are due to all those who have contributed to the current volume and the IOI conference. A lot of work is required, not only to the writing of the papers, but also to an extended period of reviewing and revising. Peer reviewing of the papers takes a significant amount of time and work, and special thanks should be given to the reviewers. Many thanks to the reviewers Ingrianni Liem, Jari Koivisto, Rein Prank, and Noa Ragonis who wrote several reports and improved quality of papers.

In particular, we would like to thank the organisational committee for IOI'2017 in Tehran and the Iranian organisation of this year's IOI for giving us the opportunity to host the conference.

Editors

Learning Trajectory of Item Response Theory Course Using Multiple Softwares

Heri RETNAWATI
Mathematics Department, Math ans Science Faculty, Universitas Negeri Yogyakarta
Yogyakarta, Indonesia
e-mail: heri_retnawati@uny.ac.id

Abstract. In a learning process with regards to parameter estimation of item response theory (IRT), which has been related to quite intricate mathematical equations, software use should be an urgent matter. The study was to describe the learning trajectory of item response theory that had been integrated to the use of multiple software and obstacles of advanced IRT course. The study was validation study-type of design research with qualitative approach. The participants in the study were the graduate school students of Educational Research and Evaluation Study Program who concentrated on Educational Measurement. The learning trajectory that had been hypothesized was based on the researcher's experience in lecturing and in identifying the necessary requirements for each chapter in the course. The data were gathered by means of observation during learning, course activities documentation and graduate school students' learning results documentation. The data were analyzed qualitatively. By formulating and testing the hypotheses of learning trajectory, the researcher found that multiple software use in course might assist the students in understanding and improving the students' skills in terms of implementing IRT although they might encounter with several obstacles.

Keywords: item response theory (IRT), software, obstacles in courses.

1. Introduction

In the activities of research, education and development in many fields, good measurement instrument is an urgent matter. Such instrument will be more urgent if the components under measurement within these activities belong to latent variables, which has been assesses through observable indicators. For this interest, the instrument that will be implemented should be analyzed in order to attain well-qualified instrument. The quality of an instrument might be seen in its validity and reliability and it is also based on the characteristics of items that compose the instrument.

Item analysis is conducted to identify the quality of each item in the instrument. This analysis might be conducted using classical test theory approach and item response theory approach. The classical approach is implemented in estimating reliability, item

difficulty and item discrimination index. Such approach is easy to implement in identifying the quality of measurement instrument; however, this approach has weakness. The weaknesses are that there is dependency between the participants' ability and the item parameters and that there is difficulty in implementing the concept of item response theory toward multiple measurement problems.

The approach of item response theory is implemented in order to overcome the weakness that that of classical test theory has. In this item response theory, the latent trait that has been measured is called as ability (Hambleton and Swaminathan, 1985; Hambleton, *et al.*, 1991). This ability is measured through the indicators of an instrument, both the test ones and the non-test ones. The relationship between ability and item parameters, such as item difficulty index, item discrimination index, pseudo guessing and alike, might be applied towards the dichotomous and polytomous data; certainly, the application should involve different equation model. The reason is that the mathematical model of this item response theory is not linear; therefore, the item parameter and ability estimations might not be directly conducted. However, through certain process, both by means of maximum likelihood and Bayes method, the item parameter and ability estimations might be directly conducted.

The skills of performing analysis to estimate item and ability parameters have been trained in universities especially in equipping university students as the researcher and psychometrical analyst candidates. These skills are trained through IRT course. In this course, the students learn multiple chapters related to item response theory or IRT, starting from classical test theory weakness, mathematical model in unidimensional IRT, parameter estimation, item characteristics curve, multidimensional IRT models and IRT implementation such as linking score, differential item functioning, cut of score and computerized adaptive testing. All IRT concepts and their implementation make use of mathematical models and several analyses, such as item parameter and ability estimations, are conducted by means of numeric approach; certainly, such approach demands software use.

Based on the course that has been taught in previous years, IRT has been conducted conventionally. In this course, the lecture has merely been explaining the theory and discussing the mathematical models without any practice of real data analysis in order to implement the IRT concepts. The software use has also been very minimum. The students' learning trajectory during the course has also been less systematic; as a result, the students have been less able to understand the inter-chapter relationship. This situation has caused the students to have less experience in analyzed the data related to instrument characteristics using IRT approach, to have difficulties when they should implement IRT in their undergraduate thesis and to have been less able in using software for assisting their analysis.

In IRT courses under measurement concentration within Educational Research and Evaluation Study Program, Graduate School of University Negeri Yogyakarta, the students come from multiple study programs (multi-entry). In one hand, this situation is beneficial because they come from both the exact study programs and the social science study programs. On the other hand, the lecturers who teach IRT become challenged because they should teach the students about multiple mathematical equations that lead to

the implemented statistics. In order to assist the item analysis of IRT, there are multiple softwares that might be used. There are several mathematical models of IRT and the name of each model depends on the item parameter that the model contains, the data that the model uses and the dimension that the model measures. The software analysis depends on the mathematical models. For the IRT using Rasch Model, the students can run QUEST (Adams and Khoo, 1993) and CONQUEST (Wu, *et al.*, 1997). For unidimensional IRT with parameter model 1, 2 and 3 on graded response model-type dichotomous data or generalized partial credit model-type dichotomous data, the students can use PARSCALE (Muraki and Block, 1997). For multidimensional IRT with dichotomous data, the students can use TESFACT (Wilson, *et al.*, 1984). For the linking score analysis, the students can use IRTEQ (Han, 2009). In order that the students might understand and use these softwares, the course should implant the concept and the strategy of application if the measurement activities are to be integrated into the software use.

Along with the development of science and technology, the integration of course implementation and software use has been an urgent matter. The integration of information and communication technology (ICT) into learning process provides multiple benefits. One of such benefits has been proposed by Ranasinghe (2009), "Integrating technology into classroom is an approach to develop better understanding of basic concepts provided it is applied appropriately." His statement has also been supported by Ochkov and Bogomolova (2015). Several experts in the past had studied the integration of learning process and ICT use like Keirns (1992), Brekke and Hogstad (2010), Thomas (2006), Hudson and Porter (2010), Aydin (2009) and use of ICT in assessment (Retnawati, 2013).

The use of computer in the learning process is influenced by several factors. Norton and Cooper (2001) states that there has been a complex interaction among the factors of "belief, pedagogical knowledge, knowledge about using technology in teaching, cultural press and perceptions of assessment" and this complex interaction is related to syllabus reformation and teacher professional development. Hudson, *et al.* (2008) state the existence of certain burden for teachers in benefitting ICT with regards to lesson plans. Another problem will be insufficient of learning plans (Hudson, *et al.*, 2008), insufficient equipment (Unal and Ozturk, 2012) and insufficient time (Salehi and Salehi, 2012; Unal and Ozturk, 2012). These factors cause the computer use in learning process has not been optimum. Budinski (2013) also states that the level of software integration in the classroom has not been satisfying.

Several strategies might help integrating ICT into learning process so that the learning process might achieve the expected results. These strategies have been proposed by Fu (2013) as follows: "more induction, orientation and training for students; an increased emphasis on the importance of instructor access and effective administration; and the expansion of podcasting and online conferencing tools." Similarly, Goktas *et al.* (2009) state that "… strategies could provide a generic approach towards enhancing this ICT integration: technology plan, in-service training, strong infrastructures, technical support and role models." In addition, professional development (Afshari *et al.*, 2009) also holds an important role in determining the successful integration of ICT into learning process.

One of the parts in lesson plan is learning trajectory design. According to Simon (1995), learning trajectory includes learning goals, learning activities and hypothetical learning process. With the presence of learning trajectory, learning becomes social aspect, continuous iterative cycle and activity that enable teachers to attain learning theories that might be adjusted to university students' ability and reasoning level (Gravemeijer *et al.,* 2003). In line with the efforts that the integration of ICT into IRT course become successful, the study is to hypothesize the learning trajectory of item response theory and its implementation by means of ICT application toward the course. The learning trajectory then will be put into experiment in order to develop the activities in IRT course.

2. Method

The study was a design research because it was to design/develop an intervention that took the form of strategy and learning material for solving the problems in education (Lidinillah, n.d.). The design research model that the researcher deployed in the study was the study validation model by Nieven et al (2006, p.152). Validation study consisted of three phases namely preparing for the experiment/preparation and design phases, design experiment and retrospective analysis (Greivemeijer and Cobb, 2006). Preparing for the experiment/preparation and design phase was conducted in order to formulate the local instructional theory by performing an analysis toward the objectives that would be achieved, such as learning objectives, and by designing the conjecture from the local instructional theory that would be developed in the form of learning trajectory. Then, design experiment had been a stage of experimental design stage that aimed to test whether the learning trajectory that had been developed would be working. Next, retrospective analysis had been a stage of analyzing the data that had been gathered in order to identify whether these data supported the learning trajectory that had been developed.

The study was conducted by designing the learning trajectory of IRT first; then, the learning trajectory would be tested in the course. The learning trajectory was designed by developing the concepts of necessary prerequisites, by analyzing the software that might be used and by planning the data that would be used in the integration of ICT toward learning process. The experiment of the learning trajectory that would be hypothesized was conducted in 2015 and 2016 toward IRT course. The participants in the course were the students from Educational Research and Evaluation Study Program who had been concentrated in measurement. In 2015 there were six students and in 2016 and there were seven students who attended the course.

The data were gathered using documentation toward learning trajectory design and observation and documentation toward learning trajectory implementation. The results of learning trajectory documentations were the results of the students' work from the practice in IRT course. These data then would be analyzed by means of Miles and Hubberman (1984) model through the steps of reduction, display and verification. The results of the analysis then would be described in order to attain the learning trajectory that would be in accordance with the design and the experiment.

3. Findings

In this section, the researcher would like to present the results of learning trajectory design activities and the results of the experiment along with the analysis.

3.1. *Designing the Learning Trajectory*

The preliminary part of the study was designing the learning trajectory. The learning trajectory was designed based on the literature review toward multiple references with regards to theory and by considering the materials that the students had studied. Such consideration was taken in relation to the prerequisite courses for IRT in the concentration of measurement under Educational Research and Evaluation Study Program.

3.1.1. *Investigating the Weakness of Classical Test Theory*

Based on the study plan of all students who concentrated on measurement, the course that they attended has been Classical Test Theory. From the information provided by these students, the weakness of this theory had been studied but it had not been investigated by using empirical data. In order that students had in-depth understanding toward the weakness of classical test theory, empirical data were necessary in order to proof the dependency of item parameters and latent ability parameters. The empirical data that had been used in the first place were the ones from participants' responses toward a test using of dichotomous scoring, the ones from participants' responses toward a test by means of polytomous scoring and the ones from participants' responses toward a questionnaire by means of polytomous scoring.

3.1.2. *Understanding the Mathematical Models in Item Response Theory with Characteristic Curve*

In item response theory, the probability of test taker to answer item correctly stated by a functional equation from the participants' ability and item parameters. This relationship might be understood by drawing a characteristic curve by implementing Rasch Model known as logistic 1 parameter model, or implementing 2 parameters logistic model, or implementing 3 parameters logistic model. The Microsoft Excel can be used to ease the calculation and in order that the students might observe the relationship pattern between the item parameters and the ability parameters. This program might also be used in drawing one characteristic curve or more.

3.1.3. *Understanding the Parameter Estimation with Excel Assistance*

After having understood the relationship between the item parameters and the ability parameters, the next step that should be taken would be estimating the ability parameters based on the already identified item parameters. The estimation might be performed using maximum likelihood approach and Microsoft Excel so that the students might un-

derstand the process step by step. Then, by deploying the raw data in the form of partici-
pants' responses toward the test with dichotomous scoring the students might estimate
the preliminary value for the item and ability parameters. With this preliminary score
parameters and maximum likelihood approach, the students might estimate the item and
ability parameters altogether. The competence of performing manual estimation pro-
vided a description toward the students in designing another program with regards to the
application of IRT.

3.1.4. *Estimating the Item Parameters and the Ability Parameters by Implementing Rasch Model with QUEST and 2PL and 3PL Model with BILOGMG*

After completing the challenge of manual estimation, the students then might practice
the estimation using computer softwares. For Rasch Model, the estimation might be
performed by QUEST program. For analysis in 2 and 3 parameters logistic model, the
estimation might be performed by BILOGMG or PARSCALE. The students might be
guided in interpreting the output of this software.

3.1.5. *Testing the IRT Assumption in Terms of Unidimension, Local Independent and Parameter Invariance*

This assumption might be studied early before the item parameter estimation and the
ability parameter estimation. However, due to the fact that the assumption testing had
made use of the item parameters and the ability parameters that had been estimated, this
step was put after the item and ability parameters estimation. The results of the estima-
tion that had been used were the item parameters estimation and the ability parameters
estimation from the data that had been split into two parts randomly based on the items
and the test participants. Then, each data would be estimated.

    The first assumption was the unidimension assumption. The assumption was applied
altogether in order to prove the second assumption, namely the local independence. The
assumption was applied in order to proof that the instrument only measure one dominant
dimension. This analysis principally calculated the eigenvalues of the inter-item correla-
tion matrix and it might be performed using SPSS software. The third assumption was
the invariance parameters that made use of the results of item and ability parameters
estimation.

3.1.6. *Understanding the Relationship of Item Parameters and Ability Parameters for Polytomous Data*

In the dichotomous data, the test participants' responses in a test or in a questionnaire
might be scored with 1 and 0. In the polytomous data, the test participants' responses
were scored 0, 1, 2, 3 and alike. Similarly in the dichotomous data, each item in a test
or in a questionnaire had item parameters and each participant had ability parameters.
The relationship between the item parameters and the ability parameters was stated in a
complex mathematical relationship and was dependent on the selected model like Rasch
Model, Partial Credit Model (PCM), and Generalized Partial Credit Model (GPCM)
and alike. The item parameters estimation might also be conducted with the assistance

of PASCALE software; the output of this software might be used in order to get an in-depth understanding toward the relationship between the item parameters and the ability parameters along with their characteristics.

3.1.7. *Understanding the Relationship Between the Item Parameters and the Ability Parameters for the Multidimension Dichotomous Data with TESTFACT*

Similarly in the unidimensional items, the relationship between the item and ability parameters for the dichotomous data in multidimensional IRT was studied. The analysis for estimating item and ability parameters was conducted by means of TESTFACT assistance and the item characteristics in this case was named as item characteristic surface. This item characteristics surface was visualized by means of MAPLE software.

3.1.8. *Implementation of IRT in Equating*

After the basic concept of IRT had been understood by the students, the next step was implementing the concept in the equaing of test set. This equating might be conducted altogether with the common items or without the common items by implementing multiple methods such as mean and mean, mean and sigma and characteristic curve. At the beginning, the students used Excel in order to understand the concept of equalization. Then, they can used IRTEQ software.

3.1.9. *Implementation of IRT to Detect DIF*

Another implementation of IRT that had been frequently used was the detection of differential item functioning (DIF). This detection was conducted by calculating the different probability in responding appropriately toward the test items of two groups under study, namely focal group and reference group. This step was taken after the item parameters and the ability parameters of each group had been estimated. The analysis might be directly conducted by means of software such as BILOGMG, PARSCALE or TESTFACT. The pattern of inter-group difference might be identified both by drawing the characteristic curve and by drawing the item characteristic surface for the items that measured multidimension.

3.1.10. *Determining the Cut of Score in a Test by Means of Bookmark Method that Had been An Implementation of IRT*

Another implementation of item response theory was the determination on the cut of score, especially bookmark method. The steps in determining the cut of score were studied and practiced by the students to gain their understanding.

3.1.11. *Estimating the Ability Parameters in CAT by Means of Simulation through Excel*

In line with the development of science and technology, the implementation of computer-based testing and computer adaptive testing (CAT) that had been adapted to the participants' ability had been an urgent matter; as a result, the estimation of partici-

pants' ability might be conducted efficiently. By combining the ability parameters and the algorithm of item determination, the students might conduct a simulation of ability estimation in IRT. This process might be conducted with the assistance of Excel that had become a bridge before the students developed the actual CAT.

The learning trajectory that had been hypothesized then would be tested into Item Response Theory course. The results displayed in each stage would be analyzed and the students would also report the obstacles that they encountered during the learning process.

3.2. *The Results of Experiment and Analysis*

The learning trajectory that had been hypothesized would be tested in Item Response Theory course. The results of each stage in the test were as follows.

3.2.1. *Identifying the Weakness of Classical Test Theory*

In order to identify the weakness of classical test theory, the students made use of the data from the national examination taken by the school that had high achievement and by the school that had low achievement. In the same time, the students also made use of the data from the documentation of Junior High School Mathematics National Examination in Indonesia. By comparing the item parameters in classical test theory between both school categories, the weakness of classical test theory regarding the dependence of item parameters, both the item difficulty index and the item discrimination index, toward the test participants' ability might be identified. Since the data that had been gathered were still raw, the students should change these data into the dichotomous data or the polytomous data.

In this step, most of the students had difficulties because they had not been used to managing huge data size. In order to change these data into dichotomous and polytomous data, the students used Excel while learning about several expressions of function that they had not understood.

3.2.2. *Understanding the Mathematical Models in IRT by Means of Characteristic Curve*

The relationship between the item parameters and the ability parameters in several models within unidimensional item response theory was investigated by means of graphic. An example of the graphic that had been drawn by the students would be displayed in Fig. 1.

The obstacle that had been encountered was that the students had not been used to the exponential function because there were about 25% of the students who had used to study and to draw function graphics. In order to draw these graphics, the lecturer's guideline was necessary for the students. Then, another obstacle that had been found was that there were 4 students who had less understanding toward the nonlinear function because the concept of linear function had been strongly embedded into their mind. By
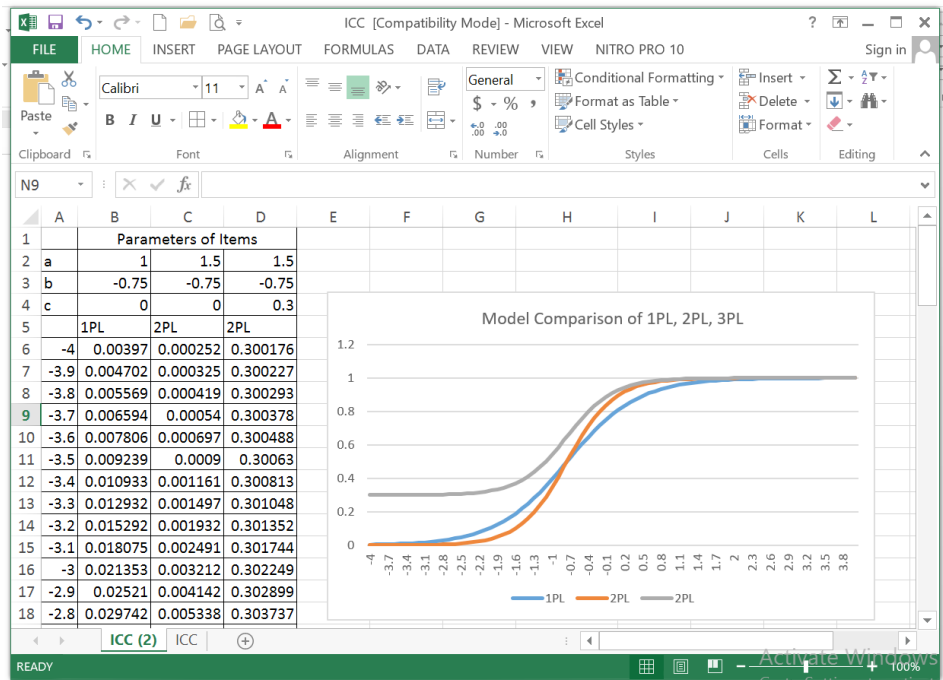
Fig. 1. Model Comparison in Unidimensional Item Response Theory.

providing the examples of nonlinear function case in the real life and the examples of other model such as the quadratic one or the cubical one, the 4 students gained better understanding. This obstacle was presented in the following statements.

> *"I have not drawn any graphics previously. Since my major is language, I have several difficulties in drawing graphics."*

> *"I just found that there is a nonlinear function. I just realized it. As far as I know, ability is always directly proportional to anything; so, my assumption is that the graphic will be straight."*

3.2.3. *Understanding the Parameter Estimation by Means of Excel Assistance*

After identifying the relationship between the item parameter and the ability parameter, the next step would be understanding the process of estimating the ability parameter based on the item parameter that had been found. The estimation method that would be implemented was maximum likelihood. At the beginning the students had difficulties in applying the likelihood equation in order to perform their estimation. In this case, the lecturer should guide the students so that they would be able to identify the parameters, both the item ones and the ability ones, that caused the likelihood to achieve maximum values. In order to ease the process, the estimation might be conducted by means of

Fig. 2. An Example of Worksheet of Ability Parameters Estimation with Maximum Likelihood.

Microsoft Excel software assistance. An example of estimation process would be presented in Fig. 2.

3.2.4. *Performing Item Parameter Estimation and Ability Parameter Estimation by Means of Software*

After performing semi-manual estimation, the students then were guided in performing the data analysis toward the test participants' responses in order to estimate the item parameters and the ability parameters by means of software. The software that would be used was dependent on the parameter models that might be used. For the analysis using Rasch Model, the students might use QUEST, 1PL model, 2PL model and 3PL model with BILOGMG. In using BILOGMG, in addition to performing item parameters estimation and ability parameters estimation, the students might also perform an analysis of model fitness both the one that used chi-square analysis and the one that used graphics. An example of output resulted the analysis toward model fitness would be presented in Fig. 3.
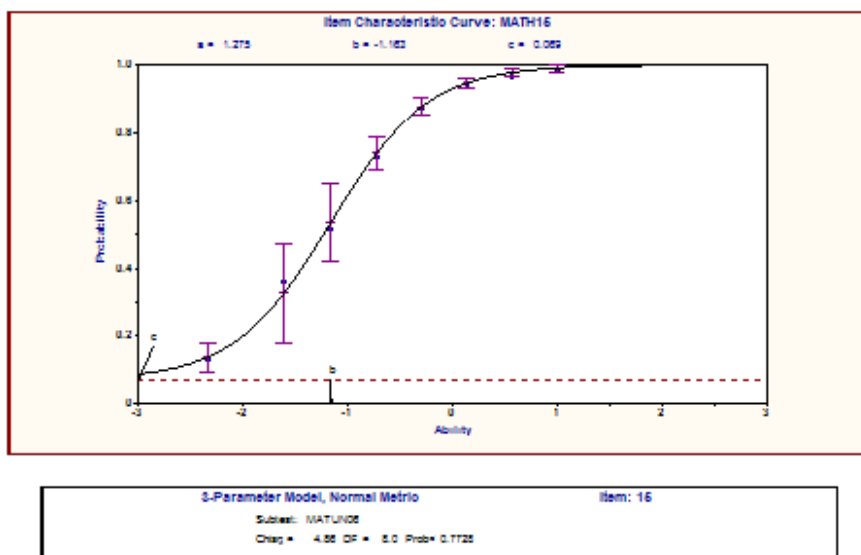
Fig. 3. An Example of Output Results from the Analysis Using 3PL Model with BILOGMG

In this analysis, the students had difficulties in conducting analysis on several stages namely preparing the data, preparing the syntax and interpreting the analysis results. In order to anticipate these problems, the lecturer should provide step-by-step tutorial toward all students with repetition. For preparing the data, the lecturer might ease this step by editing the data in the Excel's worksheet and saving these data in the *.prn format. The students' difficulties might be found in the following statement.

> *"I have difficulties in inputting the data, understanding the data and compiling the syntax analysis."*

> *"Well, this is it: reading the output results and interpreting them become my problems."*

3.2.5. *Testing the IRT Assumptions in Terms of Unidimension, Local Independence and Parameter Invariance*

The IRT assumptions in terms of unidimension were analyzed by means of exploratory factor analysis; the exploratory factor analysis was conducted by measuring the scree-plot of eigenvalues. Based on the scree-plot, the students might identify the content of a dimension that had been measured by a measurement instrument and in the same time the students might prove the local independent. The parameter invariance was proven by splitting an instrument set into two parts based on the ability; then, the students analyzed each part in order to estimate the ability parameters. Next, the students would pair the ability parameters to each participant in order to create the scatterplot. Similarly, the test participants' responses were split in order to estimate the item parameters. The results

of estimation were paired for each item in order to create the scatterplot. Proximity between the points in the scatterplot and the x = y line was used in order to prove the parameter invariance.

The assumption test was conducted by multiple software. As an example, the exploratory factor analysis might be conducted with the assistance of SPSS software and item and ability parameters estimation might be conducted with the assistance of BILOGMG software; the graphics of these analyses might be drawn with the assistance of Microsoft Excel software. Multiple software that might be implemented provided peculiar challenges for the students because they demanded adjustment and the adjustment took a relatively long time. An example of the analysis results would be presented in Fig. 4 and Fig. 5.
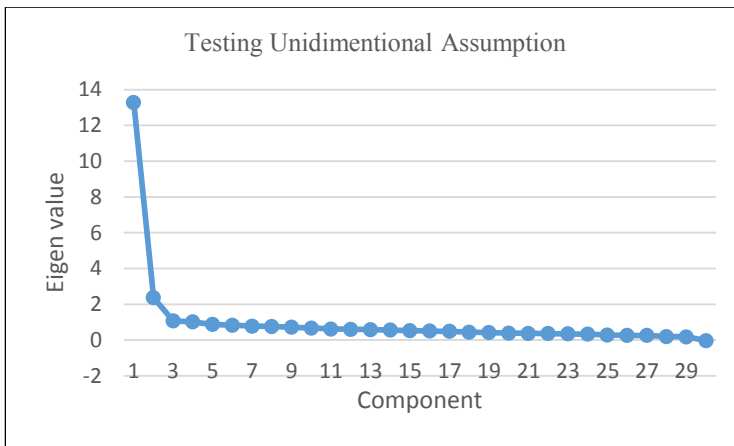
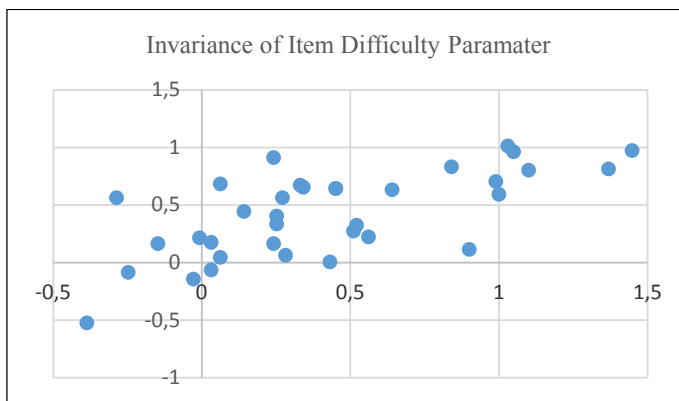Fig. 4. An Example of Scree-Plot for Proving the Unidimension

Fig. 5. An Example of Scatterplot for Proving the Parameter Invariance

3.2.6. *Understanding the Relationship between the Item and Ability Parameters for the Polytomous Data*

The students also used the similar stages in the dichotomous data for the analysis toward the polytomous data. The stages are understanding the relationship between the item parameters and the ability parameters, drawing the item characteristic curve (also known as the category response function) and estimating parameters). The parameter estimation was conducted using QUEST software for the partial credit model (PCM) and PARSCALE for the partial credit model (PCM), graded response model (GRM) and generalized partial credit model (GPCM). An example of analysis output for the item parameters would be presented in Fig. 6, for the category response function would be presented in Fig. 7 and for the value of information function and measurement errors would be presented in Fig. 8.

The difficulties that the students had in conducted the polytomous data analysis were related to the complex mathematical model and the interpretation on the output of analysis results. However, because they had experiences in analyzing the dichotomous data, it had been easier for the students to arrange the analysis syntax. The students' opinions about this stage would be presented as follows.
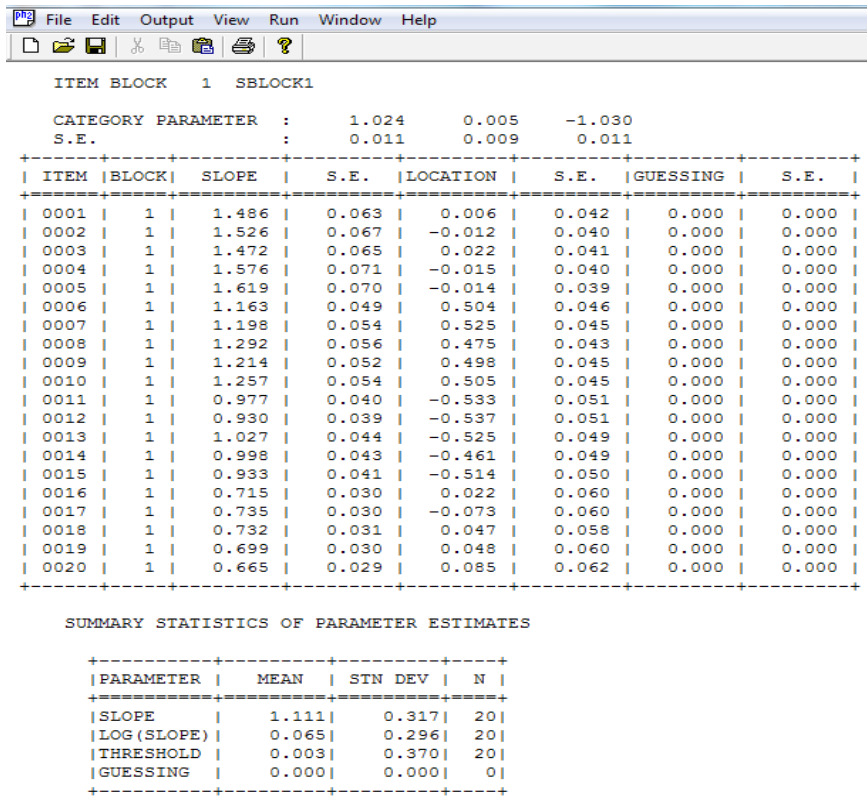


```
      File  Edit  Output  View  Run  Window  Help

      ITEM BLOCK    1   SBLOCK1

      CATEGORY PARAMETER   :      1.024       0.005     -1.030
      S.E.                 :      0.011       0.009      0.011
  +------+-----+---------+---------+---------+---------+---------+---------+
  | ITEM |BLOCK|  SLOPE  |  S.E.   |LOCATION |  S.E.   |GUESSING |  S.E.   |
  +======+=====+=========+=========+=========+=========+=========+=========+
  | 0001 |  1  |  1.486  |  0.063  |  0.006  |  0.042  |  0.000  |  0.000  |
  | 0002 |  1  |  1.526  |  0.067  | -0.012  |  0.040  |  0.000  |  0.000  |
  | 0003 |  1  |  1.472  |  0.065  |  0.022  |  0.041  |  0.000  |  0.000  |
  | 0004 |  1  |  1.576  |  0.071  | -0.015  |  0.040  |  0.000  |  0.000  |
  | 0005 |  1  |  1.619  |  0.070  | -0.014  |  0.039  |  0.000  |  0.000  |
  | 0006 |  1  |  1.163  |  0.049  |  0.504  |  0.046  |  0.000  |  0.000  |
  | 0007 |  1  |  1.198  |  0.054  |  0.525  |  0.045  |  0.000  |  0.000  |
  | 0008 |  1  |  1.292  |  0.056  |  0.475  |  0.043  |  0.000  |  0.000  |
  | 0009 |  1  |  1.214  |  0.052  |  0.498  |  0.045  |  0.000  |  0.000  |
  | 0010 |  1  |  1.257  |  0.054  |  0.505  |  0.045  |  0.000  |  0.000  |
  | 0011 |  1  |  0.977  |  0.040  | -0.533  |  0.051  |  0.000  |  0.000  |
  | 0012 |  1  |  0.930  |  0.039  | -0.537  |  0.051  |  0.000  |  0.000  |
  | 0013 |  1  |  1.027  |  0.044  | -0.525  |  0.049  |  0.000  |  0.000  |
  | 0014 |  1  |  0.998  |  0.043  | -0.461  |  0.049  |  0.000  |  0.000  |
  | 0015 |  1  |  0.933  |  0.041  | -0.514  |  0.050  |  0.000  |  0.000  |
  | 0016 |  1  |  0.715  |  0.030  |  0.022  |  0.060  |  0.000  |  0.000  |
  | 0017 |  1  |  0.735  |  0.030  | -0.073  |  0.060  |  0.000  |  0.000  |
  | 0018 |  1  |  0.732  |  0.031  |  0.047  |  0.058  |  0.000  |  0.000  |
  | 0019 |  1  |  0.699  |  0.030  |  0.048  |  0.060  |  0.000  |  0.000  |
  | 0020 |  1  |  0.665  |  0.029  |  0.085  |  0.062  |  0.000  |  0.000  |
  +------+-----+---------+---------+---------+---------+---------+---------+

      SUMMARY STATISTICS OF PARAMETER ESTIMATES

      +----------+---------+---------+----+
      |PARAMETER |  MEAN   | STN DEV | N  |
      +==========+=========+=========+====+
      |SLOPE     |  1.111| |  0.317| | 20|
      |LOG(SLOPE)|  0.065| |  0.296| | 20|
      |THRESHOLD |  0.003| |  0.370| | 20|
      |GUESSING  |  0.000| |  0.000| |  0|
      +----------+---------+---------+----+
```

Fig. 6. An Example of Output from the Results of Item Parameters Analysis Using
PARSCALE.

"The data analysis toward the polytomous scoring turns out to be more complicated than the one toward the dichotomous scoring, whereas the data analysis toward the dichotomous scoring itself has been difficult."

"The analysis syntax of the polytomous data is similar to that of the dichotomous data; this is pretty helpful."
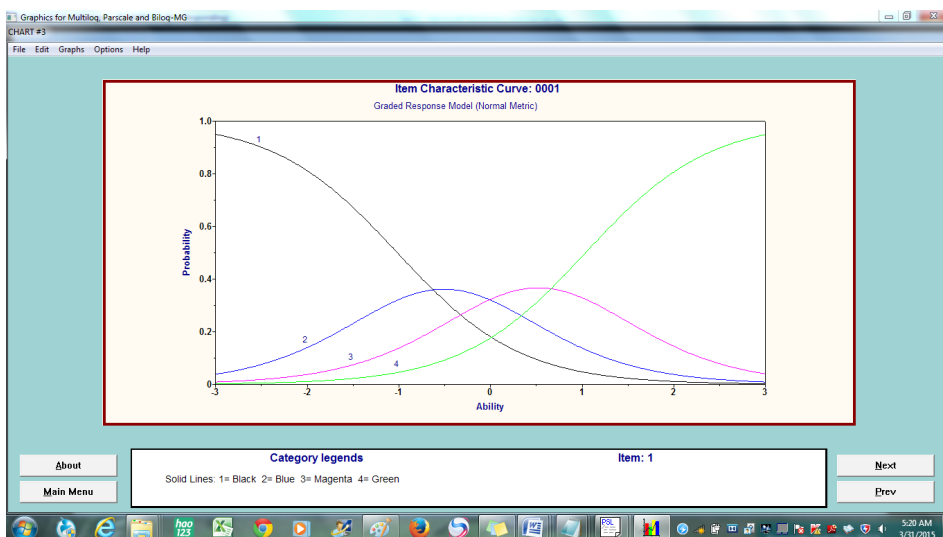

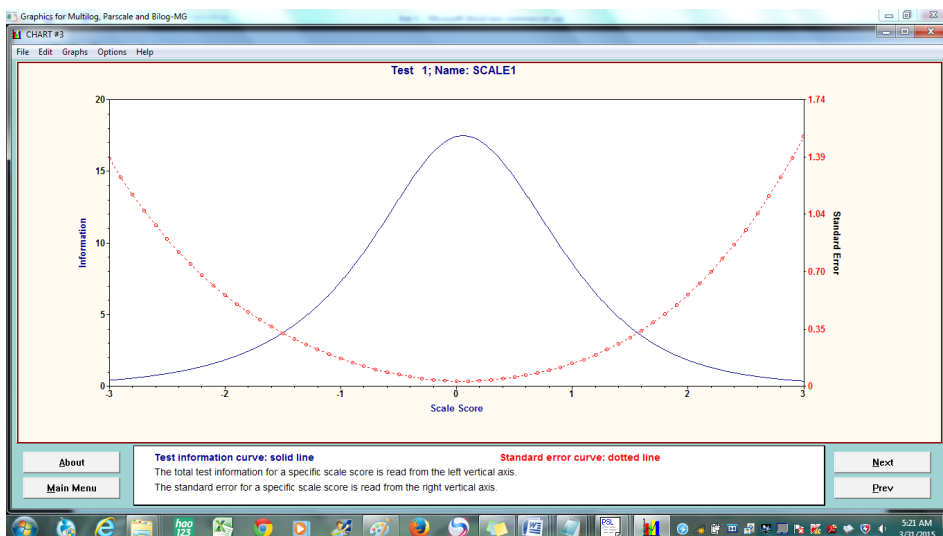Fig. 7. An Example of Output from the Item Category Response Function with GRM-Type.

Fig. 8. An Example of Output from the Value of Item Information Function and Measurement Errors.

3.2.7. *Understanding the Relationship between the Item Parameters and the Ability Parameters for the Multidimensional Dichotomous Data*

If the assumption of unidimension for IRT in the dichotomous data analysis was not prevailing, then the students would conduct an alternative analysis using multidimensional IRT. First, the researcher introduced the mathematical models of multidimensional IRT; then, the researcher would introduce the relationship between the item parameters and ability parameters. The item characteristics surface would be described with the assistance of MAPLE software. The analysis for estimating the item parameters and ability parameters would be conducted by means of TESTFACT software. An example of item characteristics surface resulted from TESTFACT software would be presented in Fig. 9.

Although at the beginning the students had difficulties in understanding the mathematical relationship in the multidimensional item response theory, they found an interesting matter. Their difficulties were caused by the complex mathematical models but the students were happy with the graphic that they made especially the items that measured 2 dimensions. The students had been accustomed to reading output from the analysis results and interpreting these results; therefore, they did not have any difficulties at all. These findings were supported by the following statements.

> *"The mathematical models of multidimensional IRT have been more complex than the last course."*

> *"Looking at the graphics, I am happy. It is easier to understand the relationship between the ability parameters and the probability of students' to respond item correctly."*

> *"The more I read the output, the easier I interpret the output."*
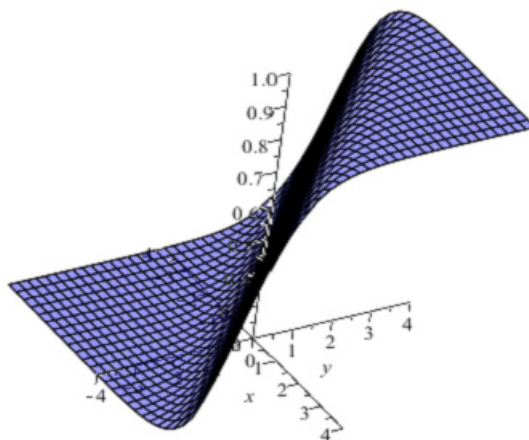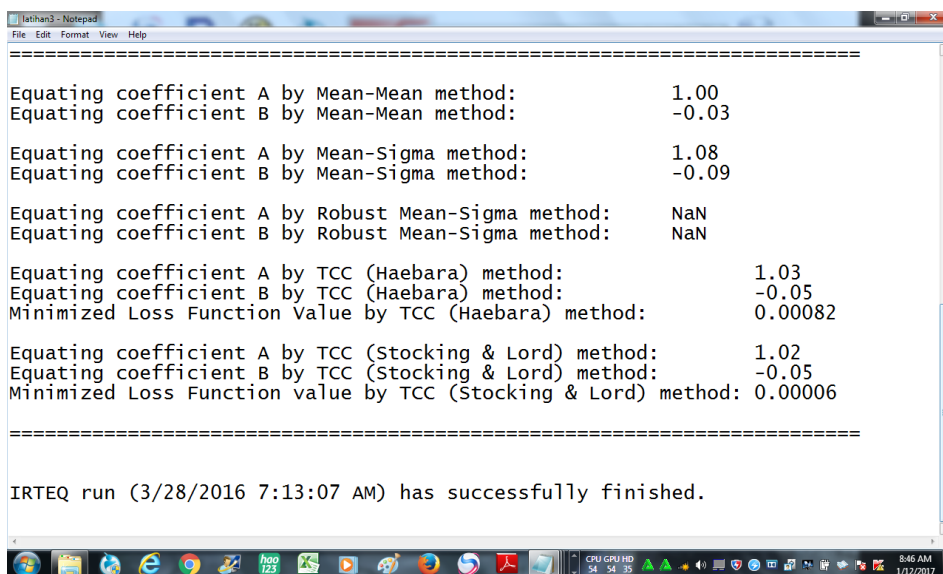
Fig. 9. An Example of Item Characteristic Surface that Measured 2 Dimensions

3.2.8. *Implementing IRT toward Equating*

By using the data of participants' responses toward several equal test sets, the students were invited to implement IRT for the linking score, both in equating manner and in concordance manner. The practice of implementing the concept of linking score was conducted semi-manually by means of Microsoft Excel software, mean and mean method, mean and sigma method and characteristic curve method (Stocking and Lord method and Haebara method). After estimating the item parameters and the ability parameters, the steps in this stage were simple yet the students had difficulties because they needed to understand the regression analysis in order to generate the prediction scores. However, after using IRTEQ program it was easier for the students to implement the concept of linking score. An example of analysis resulted by IRTEQ program would be presented in Fig. 10.

3.2.9. *Detecting Differential Item Functioning*

To detect DIF, the students categorized first the responses of focal group participants and reference group participants; then, they analyzed the item and ability parameters of both groups. The item parameter number-n for the focal group and the reference group was used in order to draw the characteristic curve so that the students would identify at which ability scale interval an item would benefit certain group. An example of results from the response data generated by the polytomous scoring would be presented in Fig. 11 and for the dichotomous scoring would be presented in Fig. 12. In this stage, the students' difficulties were more related to the technical aspects in terms of mathematical expressions in drawing graphics.

Fig. 10. An Example of Analysis Resulted by IRTEQ Program for Generating the Prediction Scores.
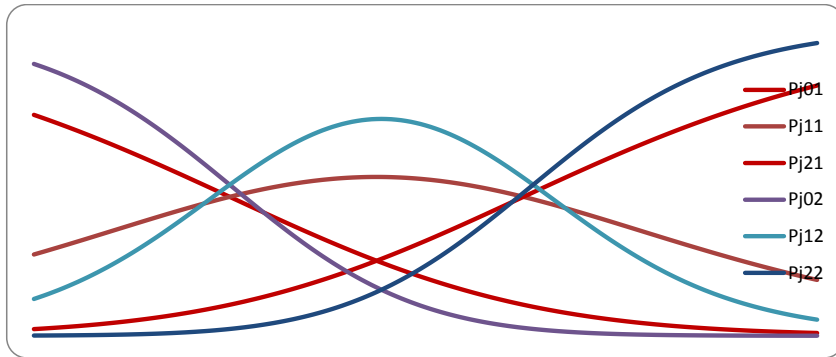
Fig. 11. An Example of Category Response Function for the Focal Group and the Reference Group.
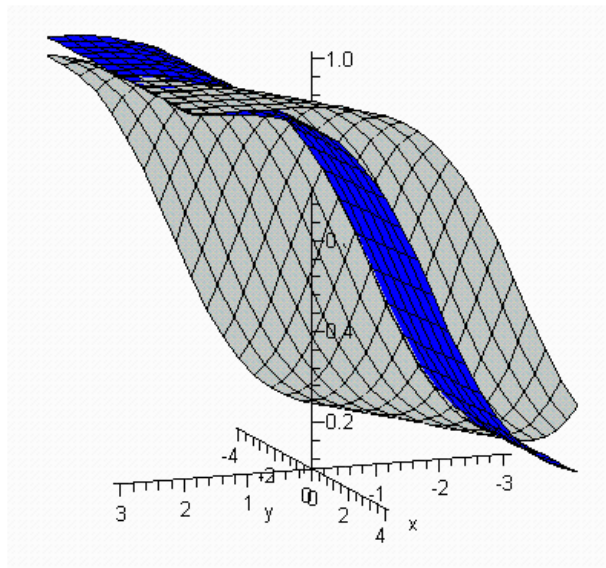
Fig. 12. An Example of Item Characteristics Surface for the Focal Group
and the Reference Group Using MAPLE.

3.2.10. *Determining the Cut of Score in a Test by Means of Bookmark Method which Had been an IRT Implementation*

By using the real data as well, the students benefitted the results of item parameters estimation in order to design the booklet and they designed the booklet by ordering the necessary minimum practice for attaining 67% chance of responding an item correctly. The booklet then was tested in order to determine the cut of score for certain lessons; in the study, the booklet was used in mathematics. In the practice, the students did not experience any significant difficulties because they also implement the insight that they gained from the course. They also had ideas to conduct studies related to the IRT implementation in the future.

3.2.11. *Estimating the Item Parameters on CAT by means of Simulation through Excel Program*

Using a set data such as item parameters, the students were invited to do simulation in CAT. They pretended as if they were arranging an item bank and then they calculated the algorithm of simple item selection like tree algorithm. Using simulation data in the form of students' response pattern, the students stimulated the test participants' ability estimation by means of one method such as maximum likelihood. In addition, the students also practiced the arrangement of stopping rule regulations within the CAT implementation.

There were several obstacles that the students encountered in studying CAT. In relation to these obstacles, there should be an in-depth understanding toward the concept of item bank and its arrangement by means of certain method like the tree algorithm. Although it had been simple, the tree algorithm should be explained by the lecturer with a real example and by involving examples from certain lessons that the students might understand. The real description regarding the example of CAT presentation should also be a topic of question for the students. However, these multiple questions had been the trigger for the students to conduct more studies in the future. This finding was supported by the following statements.

> "After learning CAT, I have many questions in my mind. Anyway, the positive aspect is that I would like to conduct research about CAT."

> "It turns out that, there have been so many ideas of research that we have, especially in instrument item characteristics and its implementation such as equating, cut of score and CAT."

4. Conclusions and Discussions

Paying attention to the results of the analysis toward each experimental stage in the learning trajectory, the researcher found that the learning trajectory of item response theory through the software use that had been hypothesized had provided the expected results. In the study, the students might interpret their understanding through the output that had been resulted from the practice, especially their understanding toward relationship between item and ability parameters. The students had also been able to use multiple software in performing real data analysis and simulated data analysis by implementing multiple concepts in item response theory. This had been a trigger for the students to perform their studies by implementing the theory.

This success was supported by several matters. First, by hypothesizing the learning trajectory, the lecturer actually had already planned and had already prepared the course, including the learning materials, the software and the supporting data. With sufficient preparation, the course might run well. This situation had been in accordance with the suggestions by Hudson *et al.* (2008) that the successful learning should be supported by a good learning plan. Second, the learning trajectory that had been planned had given attention toward the necessary prerequisite materials. The materials that should be stud-

ied were supported by the materials that had previously been mastered by the students. Third, the instructor had been very encouraging; the instructor had trained the integration of ICT into the learning process step by step along with its application toward the related problems and the use of existing software; as a result, it would be easier for the students to understand the materials of item response theory especially the materials that had been related to the mathematical equations and the courses that made use of multiple complex software. The third aspect here had been in accordance with the results of a study by Ranasinghe (2009).

The students encountered multiple obstacles during while they were integrating the technology into the learning process. Their obstacles within the learning process were coming from multiple study programs (multi entries), having not been accustomed to using Microsoft Excel, having not been accustomed to arranging analysis syntax, having to deal with complex mathematical equations and having to use multiple software. The strategies that had been taken in order to overcome these obstacles were designing handouts for the students and training the students to have step-by-step understanding toward the concept in details. It took a long time to implement these strategies. These findings, by the way, had been in accordance to the results of several studies (Salehi and Salehi, 2012; Unal and Ozturk, 2012). The students also had difficulties in using multiple software. Therefore, using single software that might cover all types of data analysis in IRT course is an urgent matter that should be pursued.

References

Adams, R.J., Khoo, S.T. (1993). *Quest: The Interactive Test Analysis System [Computer Program]*. Hawthorn: Australian Council for Educational Research.

Afshari, M., Abu Bakar, K., Luan, W., S., Abu Samah, B., Fooi, F.,S. (2009). Factors affecting teachers' use of information and communication technology**.** *International Journal of Instruction,* 2(1), 77–104.

Aydin, E. (2005). The use of computers in mathematics education: a paradigm shift from "computer assisted instruction" towards "student programming". *The Turkish Online Journal of Educational Technology – TOJET*, 4(2), 27–34.

Brekke, M., Hogstad, P.H. (2010). New teaching methods – Using computer technology in physics, mathematics and computer science. *International Journal of Digital Society* (*IJDS*), 1(1), 17–24.

Budinski, N. (2013). A survey on use of computers in mathematical education in Serbia. *The Teaching of Mathematics*, 16(1), 42–46.

Du Toit, M. (2003). *IRT from SSi: BILOG-MG, MULTILOG, PARSCALE, TESTFACT*. Lincolnwood: SSi.

Fu, J.S. (2013). ICT in education: A critical literature review and its implications. *International Journal of Education and Development using Information and Communication Technology (IJEDICT)*, 9(1), 112–125.

Ghavifekr, S., Rosdy, W.A.W. (2015). Teaching and learning with technology: Effectiveness of ICT integration in schools. *International Journal of Research in Edution and Science (IJRES)*, 1(2), 175–191.

Goktas, Y., Yildirim, S., Yildirim, Z. (2009). Main barriers and possible enablers of ICTs integration into pre-service teacher education programs. *Educational Technology & Society*, 12(1), 193–204.

Gravemeijer, K., Cobb, P. (2006). Design research from a learning perspective. In: van den Akker, J., Gravemeijer, K., McKenney., S., Nieveen, N. (Eds.), *Educational Design Research*. New York: Routledge.

Gravemeijer, K., Bowers, J., Stephan, M. (2003). A hypothetical learning trajectory on measurement and flexible arithmetic. *Journal for Research in Mathematics Education. Monograph,* 12, 51–66.

Hambleton, R.K., Swaminathan, H. (1985). *Item Response Theory*. Boston, MA: Kluwer Inc.

Hambleton, R.K., Swaminathan, H., Rogers, H.J. (1991). *Fundamental of Item Response Theory*. Newbury Park, CA: Sage Publication Inc.

Han, K.T. (2009). IRTEQ: Windows application that implements IRT scaling and equating [Computer pro-

gram]. *Applied Psychological Measurement*, 33(6), 491–493.

Hudson, R., Porter, A. (2010). ICT use to improve mathematics learning in secondary schools. *Paper in ACEC2010: Digital Diversity Conference*, 6–9 April 2010, Melbourne, Australia.

Hudson, R., Porter, A.L., Nelson, M.I. (2008). Barriers to using ICT in mathematics teaching: issues in methodology. In: Luca, J., Weipl, E. (Eds.), *ed-Media 2008 World Conference on Educational Multimedia, Hypermedia and Telecommunications*. Chesapeake, VA: Association for the Advancement of Computing in Education, 5765–5776.

Keirns, J.L. (1992). Does computer coursework transfer into teaching practice? A follow-up study of teachers in a computer course. *Journal of Computing in Teacher Education*, 8(4).

Lidinillah, D.A.M. (n.d.). *Educational Design Research: a Theoretical Framework for Action*. Available at www.file.upi.edu/Direktori/

Miles, M.B., Hubberman, A.M. (1994). *Qualitative Data Anaysis*. California: SAGE Publications.

Mislevy, R.J., Bock, R.D. (1990). *BILOG 3: Item Analysis & Test Scoring with Binary Logistic Models [Computer Program]*. Moorseville: Scientific Sofware Inc.

Muraki, E., Bock, R.D. (1997). *Parscale 3: IRT Based Test Scoring and Item Analysis for Graded Items and Rating Scales [Computer Program]*. Chicago: Scintific Software Inc.

Nieveen, N., McKenney, S., van den Akker (2006). Educational design research. In: van den Akker, J., Gravemeijer, K., McKenney., S., Nieveen, N. (Eds.), *Educational Design Research.* New York: Routledge.

Norton, S., Cooper, T. (2001). Factors influencing computer use in mathematics teaching in secondary schools. *Paper in 24th Annual MERGA Conference*, July 2001, Sydney, Australia, 386–393.

Ochkov, V.F., Bogomolova, E.P. (2015). Teaching mathematics with mathematical software, *Journal of Humanistic Mathematics*, 5(1), 265–285. DOI: 10.5642/jhummath.201501.15. Available at: http://scholarship.claremont.edu/jhm/ vol5/iss1/15

Ranasinghe, A.I., Leisher, D. (2009). The benefit of integrating technology into the classroom. *International Mathematical Forum*, 4(40), 1955–1961.

Retnawati, H. (2015). The comparison of accuracy scores on the paper and pencil testing vs. computer based testing. *The Turkish Online Journal of Educational Technology* (TOJET) 14(4), 135–142.

Salehi, H., Salehi, Z. (2012). Challenges for using ICT in education: Teachers' insights. *International Journal of e-Education, e-Business, e-Management and e-Learning*, 2(1), 40–43.

Simon, M.A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education,* 26, 114–145.

Stultz, S.,L. (2013). The effectiveness of computer-assisted instruction for teaching mathematics to students with specific learning disability. *The Journal of Special Education Apprenticeship*, 2(2), Article 7. Available at: http://scholarworks.lib.csusb.edu/josea/vol2/iss2/7

Thomas, O.J. (2006). Teachers using computers in mathematics: A longitudinal study. In: Novotná, J., Moraová, H., Krátká, M., Stehlíková, N. (Eds.), *Proceedings 30 Conference of the International Group for the Psychology of Mathematics Education*, 5, 265–272. Prague: PME5 – 265.

Unal, S., Ozturk, I.H. (2012). Barriers to ITC integration into teachers' classroom practices: Lessons from a case study on social studies teachers in Turkey. *World Applied Sciences Journal*, 18(7), 939–944, DOI: 10.5829/idosi.wasj.2012.18.07.1243

Wilson, D., Wood, R., Gibbons, R. (1984). *TESTFACT: Test Scoring and Fullinformation Item Factor Analysis* [*Computer Program*]. Mooresville, IL: SSi.

Wu**, *M.L., Adams, R.J., Wilson, M.R. (1997). ConQuest: Multi-Aspect Test Software [Computer program].* Camberwell: Australian Council for Educational Research.

H. Retnawati is Associate Professor in Mathematics Department, Mathematics and Science Faculty, Universitas Negeri Yogyakarta, Indonesia, as researcher and lecturer. Her research interest are teaching methods and assessment in mathematics education, measurement in education, and psychometric.

About Journal and Instructions to Authors

OLYMPIADS IN INFORMATICS is a peer-reviewed scholarly journal that provides an international forum for presenting research and developments in the specific scope of teaching and learning informatics through olympiads and other competitions. The journal is focused on the research and practice of professionals who are working in the field of teaching informatics to talented student. OLYMPIADS IN INFORMATICS is published annually (in the summer).

The journal consists of two sections: the main part is devoted to research papers and only original high-quality scientific papers are accepted; the second section is for countries reports on national olympiads or contests, book reviews, comments on tasks solutions and other initiatives in connection with teaching informatics in schools.

The journal is closely connected to the scientific conference annually organized during the International Olympiad in Informatics (IOI).

Abstracting/Indexing

OLYMPIADS IN INFORMATICS is abstracted/indexed by:

- Cabell Publishing
- Central and Eastern European Online Library (CEEOL)
- EBSCO
- Educational Research Abstracts (ERA)
- ERIC
- INSPEC
- SCOPUS – Elsevier Bibliographic Databases

Submission of Manuscripts

All research papers submitted for publication in this journal must contain original unpublished work and must not have been submitted for publication elsewhere. Any manuscript which does not conform to the requirements will be returned.

The journal language is English. No formal limit is placed on the length of a paper, but the editors may recommend the shortening of a long paper.

Each paper submitted for the journal should be prepared according to the following structure:

- concise and informative title
- full names and affiliations of all authors, including e-mail addresses
- informative abstract of 70–150 words

- list of relevant keywords
- full text of the paper
- list of references
- biographic information about the author(s) including photography

All illustrations should be numbered consecutively and supplied with captions. They must fit on a 124 × 194 mm sheet of paper, including the title.

The references cited in the text should be indicated in brackets:
- for one author – (Johnson, 1999)
- for two authors – (Johnson and Peterson, 2002)
- for three or more authors – (Johnson *et al.*, 2002)
- the page number can be indicated as (Hubwieser, 2001, p. 25)

The list of references should be presented at the end of the paper in alphabetic order. Papers by the same author(s) in the same year should be distinguished by the letters a, b, etc. Only Latin characters should be used in references.

Please adhere closely to the following format in the list of references:

For books:

Hubwieser, P. (2001). *Didaktik der Informatik*. Springer-Verlag, Berlin.

Schwartz, J.E., Beichner, R.J. (1999). *Essentials of Educational Technology*. Allyn and Bacon, Boston.

For contribution to collective works:

Batissta, M.T., Clements, D.H. (2000). Mathematics curriculum development as a scientific endeavor. In: Kelly, A.E., Lesh, R.A. (Eds.), *Handbook of Research Design in Mathematics and Science Education*. Lawrence Erlbaum Associates Pub., London, 737–760.

Plomp, T., Reinen, I.J. (1996). Computer literacy. In: Plomp, T., Ely, A.D. (Eds.), *International Encyclopedia for Educational Technology*. Pergamon Press, London, 626–630.

For journal papers:

McCormick, R. (1992). Curriculum development and new information technology. *Journal of Information Technology for Teacher Education*, 1(1), 23–49.
`http://rice.edn.deakin.edu.au/archives/JITTE/j113.htm`

Burton, B.A. (2010). Encouraging algorithmic thinking without a computer. *Olympiads in Informatics*, 4, 3–14.

For documents on Internet:

International Olympiads in Informatics (2008).
`http://www.IOInformatics.org/`

Hassinen, P., Elomaa, J., Ronkko, J., Halme, J., Hodju, P. (1999). *Neural Networks Tool – Nenet (Version 1.1)*.
`http://koti.mbnet.fi/~phodju/nenet/Nenet/General.html`

Authors must submit electronic versions of manuscripts in PDF to the editors. The manuscripts should conform all the requirements above.

If a paper is accepted for publication, the authors will be asked for a computerprocessed text of the final version of the paper, supplemented with illustrations and tables, prepared as a Microsoft Word or LaTeX document. The illustrations are to be presented in TIF, WMF, BMP, PCX or PNG formats (the resolution of point graphics pictures is 300 dots per inch).

Contacts for communication

Valentina Dagienė
Vilnius University Institute of Mathematics and Informatics
Akademijos str. 4, LT-08663 Vilnius, Lithuania
Phone: +370 5 2109 732
Fax: +370 52 729 209
E-mail: valentina.dagiene@mii.vu.lt

Internet Address

All the information about the journal can be found at:

`http://ioinformatics.org/oi_index.shtml`

Olympiads in Informatics

Volume 11, 2017

Olympiads in Informatics

Volume 11, 2017

9 771822 773007