Proceedings of the World Congress on Engineering 2007 Vol I
WCE 2007, July 2 - 4, 2007, London, U.K.

Assessment of Complex Abilities in Engineering Students using Structural Equation Models and Multidimensional Item Response Theory: Preliminary Results

Carlos Villegas-Quezada, Rebeca de los Santos, and C. Rodrigo Villegas

***Abstract*— Nowadays, several abilities in engineering students are required, one of them is the ICT (Information and Communication Technologies) literacy proficiency. The ICT literacy proficiency is defined as the ability to use digital technology, communication tools, and networks appropriately to solve information problems in order to function in an information society. This includes the ability to use technology as a tool to research, organize, evaluate, and communicate information. A computer system to diagnostic ICT literacy proficiency in engineering students is developing. This system includes interactive simulations which emulates the features of real software products (word processing, electronic sheet, internet search engine, presentation system and Chat). This system also includes monitors which follows the test taker's progress through the simulation to determine if the students answered a question correctly. In this paper is presented the model for validating the cognitive structure and scoring the ICT literacy proficiency test in a computer-based system. The model concerns the confluence of measurement models, statistical inference, and cognitive diagnosis in the design and analysis of an assessment model of ICT literacy proficiency. The system assessment is a comprehensive test of ICT proficiency specifically designed for engineering students. The system uses real-time, scenario-based tasks (not multiple-choice questions) to assess five ICT skills: obtaining information, managing information, evaluating, creating and communicating. When the purpose of an assessment includes what kind of knowledge, skill or ability we want to make inferences about, we need new assessment paradigms. Taking some ideas from Evidence Centered Assessment Design and the Rule Space approach, we proposed a Model to evaluate a cognitive structure and identifying several evidences that represent the knowledge, skills, and abilities we wish to assess**

The proposed method of cognitive structure validation makes the assumption that for each test item a continuous latent variable exist, which represents the ability required for correct solution of the ICT test item. The proposed model has three parts. The first one is the cognitive evaluation of skills structure, using structural equation model (with LISREL). The second one is a series of Evidence Rules that describe how to identify and evaluate essential features of the Work Product in information technologies. The third one is a statistical model based in multidimensional item response theory (MIRT) for scoring the ICT test. The MIRT considers the relationship between the unobservable variables (ICT skills), conceptualized as constructs, and the probability of the examinee to correctly answer any particular item of the ICT literacy test. The model can be extended to the assessment of complex abilities in other areas of the engineering teaching, such as: computer programming, mathematics, civil engineering, industrial engineering, etc.

***Index Terms*—Engineering abilities assessment, ICT literacy assessment, Multidimensional assessment, SEM and MRCML techniques.**

Manuscript received March 30, 2007. This work was supported in part by FICSAC (Fomento de Investigación y Cultura Superior, A.C.) and Departamento de Ingenierías – UIA, México, D.F.

Carlos Villegas-Quezada is with the Engineering Department, Universidad Iberoamericana, Campus Santa Fé, México, D.F. 05330 (phone: 55-59504000 Ext. 4605; fax: 55-59504064; e-mail: carlos.villegas@ uia.mx).

Rebeca de los Santos, was with Universidad Anáhuac, México, D.F. She is now Freelance consulting in Education Sciences and E-Learning (e-mail: rebesantos@gmail.com).

C. Rodrigo Villegas is student at The Edron Academy, México, D.F. (e-mail: realmadrid2010@gmail.com).

I. Introduction

Recent decades have witnessed advances in the cognitive, psychometric, technological tools, concepts, and theories that are germane to educational assessment. Today, the need to promote abilities learning in the different educational levels is necessary. Also, we need to use new tests and techniques focused to abilities assessment.

This paper describes a framework for assessment of engineering students. Specifically an ICT literacy proficiency test. However, it can be extended to different engineering areas. The ETS defines ICT literacy proficiency as the ability to use digital technology; communication tools and networks appropriately to solve information problems in order to function in an information society. This includes the ability to use technology as a tool to research, organize, evaluate and communicate information [5]. This definition emphasizes that ICT Literacy is predominantly a cognitive activity, with a fundamental technical capability required in order to interact with technological tools to execute cognitive strategies in information retrieval, use, and dissemination.

The present research forms a part of a project to elaborate a computer system for ICT proficiencies assessment in engineering students, by using a standardized test. Specifically, the present work is focused on the validation of the cognitive model of abilities and the assessment of every student, from the answers that they provide on the ICT test. The majority of the assessments in ICT are based on applications of multiple choice test, orientated to the demonstration of knowledge more than to the demonstration of the proficiencies in ICT.

The system reported in the present paper is oriented to the development of standardized computer-based-test to evaluate abilities in ICT literacy. The students should know how to use the basic software: word processor, spreadsheet, mail and internet. And have six ICT proficiencies: access, manage, integrate, evaluate, create, and communicate. The technical elements of the computational tools and the six proficiencies must be used in an adequate way to solve the basic ICT problems.

II. THE ICT LITERACY AND THE DESIGN OF ASSESSMENT MODEL

The advances in education sciences, cognitive sciences, measurement theory, and computer sciences, among others, allow to raise powerful models and tools that have been emerging in the last years, to support the development of evaluations orientated to measure complex abilities. For the design of the test, we took some elements of the Evidence Centered Design (ECD) which is a methodology for the design of evaluation that emphasizes a logical and explicit representation of a chain of evidences based on reasoning [12].

The ECD framework is composed by four components [2], [13]: a student model that specifies the variables in the terms of which we wish to characterize students; task models, schemas for ways to get the data that provides evidence about them; the scoring component of the evidence model, contains procedures for extracting the salient features of the student's performances in task situations (observable variables); the measurement component, which contains machinery for updating beliefs about student-model variables in light of this information. The two ultimate components are the object of this paper. Also, the ICT test for engineering students assume the existence of six proficiencies:

1) Access: the ability to collect and/or retrieve information in digital environments. The ability to identify likely digital information sources and to get the information from these sources.
2) Manage: The ability to apply an existing outline or classification scheme for digital information. The ability to identify preexisting organization schemes, select appropriate schemes for the current usage, and to apply the scheme.
3) Integrate: The ability to interpret and represent digital information. The ability to use ICT tools to synthesize, summarize, compare, and contrast information from multiple digital sources.
4) Evaluate: The ability to determine the degree in which the digital information satisfies the needs of the task in ICT test. The ability to judge the relevant, authority, currency or accuracy of digital information and resources.
5) Create: The ability to generate information by adapting, applying, designing, or inventing information in the ICT test.
6) Communicate: The ability to communicate information properly in context in the ICT test. The ability to adapt electronic information for a particular audience and to communicate knowledge in the appropriate venue.

The six proficiencies (latent variables) constitute the construct or principal latent variable of ICT Literacy Test for engineering students.

III. VALIDATION OF COGNITIVE STRUCTURES WITH STRUCTURAL EQUATION MODELS

The ICT literacy test for engineering students consists of several tasks that provide evidences through observable items. These items have interrelationship with one or more of the six proficiencies of ICT. The cognitive structure of given tests is defined by the set of cognitive operations and process, and their relationships required for obtaining correct answers on its items. Therefore, to represent the cognitive structure that sublies to every set of tasks that compose the ICT test, it is assumed the existence of a cognitive multidimensional structure.

Some ideas of the Rule Space methodology proposed by Tatsuoka are used [16]. The Rule Space method represents an approach for generating diagnostic scores, characterized by, first identifying a vector of attributes (knowledge, skills, and proficiencies being tested), and then defining a Matrix Q which shows which attributes are used in which items. Also, it is proposed that the cognitive structure of a given test is defined by the set of cognitive operations and processes, as well as their relationships, required for obtaining correct answers on its items.

From the Q-Matrix, a model of structural equations (SEM) is formed which represents the cognitive multidimensional model evidences-abilities (Figure 1). This model is validated with LISREL software. The presently proposed method of cognitive structure validation makes the assumption that for each test item exists one or several continuous latent variables or constructs, which represents the ability or abilities required for correct solution of the item [3]. Students solve an item correctly if they present a sufficient proficiency. That is, if their ability relevant to the item exceeds or equals a certain minimal level required for its correct solution [7].

The suitable approach of the cognitive model of the proficiencies that will be evaluated, it is of vital importance. With the intention of verifying that the underlying theory to the latent model can predict the phenomenon observed in the information of test. For the latent model that represents the skills in ICT, there have been used ideas of the method of Rule Space [16] and ECD [2]. Later, the latent model that appeared is fitted using the software LISREL [8]. One of the preliminary diagrams (SEM), the most complex, used for the adjustment of the cognitive structure of ICT literacy, is showed in Figure 1 with path-diagram.

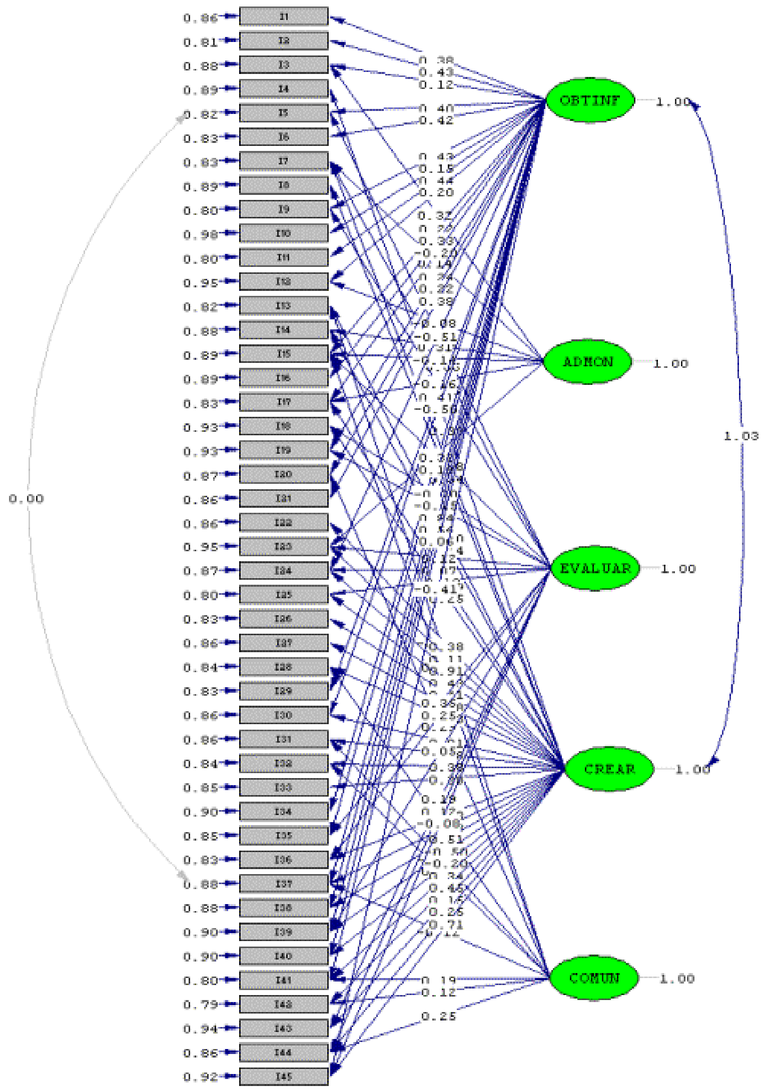

Figure 1. Preliminary SEM model used for the adjustment of cognitive structure

In this diagram, each item is represented by a rectangle, and the circles denote the several latent proficiencies required for its correct solution. This diagram represents five abilities and 45 items, in the assessment system for engineering students.

With LISREL software, the latent models are fitted, validating the several cognitive structures that compose the ICT literacy test. Using LISREL it is realized the evaluation of the diverse cognitive models items-proficiencies for the different tasks that compose the exercises of the test. In this evaluation there were deleted, added and changed items causing a change with the relation and the proposed proficiencies. Therefore it is obtained the cognitive final model that is forming the base of items to use in the test of ICT literacy.

IV ASSESSMENT OF THE PROFICIENCY OF ENGINEERING STUDENTS

For the assessment of each student, we use a scoring model based in a rule-based approach and multidimensional item response theory (MIRT) [1], [19]. The proficiency scoring model providing conceptual relationships between observations (items) and proficiency estimates (latent variables), based in each one of the path diagrams for each task in the system, and that were fitted by SEM using LISREL.

As well, the proficiency scoring model establishes the targets of the assessment and, the latent variables that the statistical model (MIRT) must be able to compute (six possible proficiencies for which tasks provide evidence items). Each task elicits performance (observable elements of behavior) that have been judged to constitute evidence. The scoring model has

ISBN:978-988-98671-5-7

three components:

1) Evidence identification: Determines what elements of the task performance constitute evidence and summarize their values.
2) Evidence accumulation: Aggregate evidence to update estimates of abilities in each proficiency.
3) Compute probabilistic proficiency: Using a multidimensional item response theory, compute the probabilistic proficiency for each student.

A student model is defined in terms of variables (θ) that represent the facets of skills or abilities that suit the purpose of the assessment, and the data (X), are values of variables that characterize aspects of the observable behavior of ICT test.

For the case of the ICT system of engineering students assessment, a multidimensional model is used. Such types of models provide a more realistic exploration of the situations of the real world, in where the performance of the student is not possible to model via a unidimensional construct (since it is postulated in the classic Item Response Theory). Specifically, the system that is postulated in the present work, uses a Multidimensional Model Random Coefficients Multinomial Logit (MRCML) proposed by Adams, Wilson and Wnag [1], to produce inferences over the skills of the students. The software used was ConQuest [20].

A. *The MRCML Model*

The latent cognitive structure (validated with SEM) was used for the scoring model. So, proficiencies, evidences, and tasks establish the characteristics of the scoring model by providing conceptual relationships between observations and proficiency estimates. For each proficiency model (cognitive structure), is established the targets of the assessment and, therefore, the latent variables that the statistical models (MIRT) must be able to accommodate.

Specifically the MRCML is used to provide a generalized solution for a family of multidimensional polytomous Rasch-based models. The MRCML family of measurement models use two matrix (A & B) to define many different measurement models. The scoring matrix (B), represents the relantionships between items (rows) and dimensions or constructs (columns). The *design matrix* (A), represents the relantionship between items (rows) and the model parameters (columns), such as item category, difficulties, step difficulties, etc.

The system uses the scoring guides to rate student performance into five ordered, qualitatively different categories, labeled zero through four. The scoring guides describe the kind of performance that can be expected from students at each of the performance levels; a score of 0 indicates an off-task or missing response; a score of 1 indicates that the performance was incorrect; a score of 2 indicates that the performance is generally correct but missing in some important things; a score of 3 indicates that the student completed a correct performance; and a score of 4 indicates excellent performance, as is showed in Figure 2.

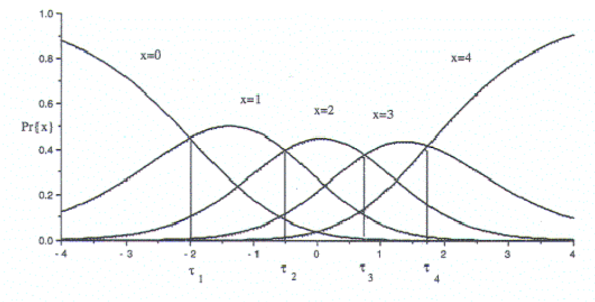

Figure 2. Model MRCML for scoring an item of ICT literacy test.

B. *Evaluation of the Scoring Model using Simulated Students*

It have been performed a preliminary number of the assessment computer-system with small groups of students. Soon, a test will be realized with a bigger sample to prove the system of ICT literacy assessment. Meanwhile, simulations have been realized to prove the cognitive models and the scoring model with MRCML.

The simulation of students' answers has been realized, considering diverse levels of complexity of the items and diverse levels of skill of the students. Simulations with 2000 students were considered, with levels of skill from very low, to very high proficiencies in ICT literacy tasks. The assessment simulation of student's proficiency was realized with the software ConQuest. In table 1, it can be observed some results of five ICT proficiencies (θ) for five students. The results are show in the IRT's traditional scale among –3.0 to +3.0 [19]. Also, the system obtains the proficiencies for the whole group of students, and the level of ability in each latent proficiency.

Table 1. Proficiencies results for some students using rule based approach and MRCML

Access (θ1)	Manage (θ2)	Evaluate (θ3)	Create (θ4)	Comunicate (θ5)
1.05713	0.02976	0.41047	0.87650	0.53681
0.30747	0.15008	0.47549	0.38121	0.63972
0.58174	0.24051	0.01968	0.45212	0.35541
0.21806	0.27324	0.10364	0.39440	0.35629
1.05262	0.27696	0.26746	1.15894	0.29790

V. CONCLUSIONS

This paper outlines the scoring model design of the ICT Literacy Assessment for Engineering Students, and Internet software assessment that measures a engineering student's proficiencies to use technology to: access, manage, create and communicate information.

This assessment is focused on the cognitive aspects of ICT literacy and it uses a simulation-based tasks (via software application) to elicit evidence of proficiencies. It uses some ideas of Evidence-Centered Design, Structural Equation Models for validate the cognitive structure, and MIRT for scoring the proficiency of each student. It is considered that the results that have been obtained with a student's small sample and with simulated students are satisfactory and accomplish with the SEM and IRT's theorical models.

The model will be applied to student's statiscally significant sample in a short period. Also, the model can be extended to the assessment of complex abilities in other areas of the engineering teaching, such as: software engineering, computer science, mathematics, civil engineering, industrial engineering, etc.

REFERENCES

[1] R. Adams, M. Wilson, and W. Wang, "The multidimensional random coefficients multinomial logit model", Applied Psychological Measurement, vol. 21, 1997, pp. 1-23.
[2] R.G. Almond, L.S. Steinberg, and R.J. Mislevy, "A four-process architecture for assessment delivery with connections to assessment design", Journal of technology, Learning and Assessment, vol. 1, no. 5, 2002.
[3] D.M. Dimitrov, "Validation of cognitive structures: A structural equation modeling approach", Multivariate Behavioral Research, vol. 38, No. 1, 2003, pp. 1-23.
[4] S.E. Embretson, "The new rules of measurement", Psychological Assessment, vol. 8, No. 4, 1996, pp. 341-349.
[5] ETS, Digital transformation: A framework for ICT literacy, Report of the international ICT literacy panel. ETS Report, 2001.
[6] ETS, Measuring College-Level information and communication technology proficiency, ETS Report, 2006.
[7] D.H. Gitomer, D. Rock, "Addressing process variables in test analysis". In N. Frederiksen, R.J. Mislevy, and I. Bejar (Eds.), Test theory for a new generation of tests (pp. 243-268), 1993, Hillsdale, NJ, Erlbaum
[8] K.G. Jöreskog, D. Sörbom, LISREL 8: User's reference guide, Chicago, IL., Scientific Software International, 1996.
[9] C.A. Kennedy, "Models and tools for drawing inferences from student work", Annual Meeting of the American Education Research Association, 2005, Montreal, Canada.
[10] C.A. Kennedy, Constructing measurement models for MRCML estimation: A primer for using the BEAR scoring engine, BEAR technical report series 2005-04-02, Berkeley, CA, Berkeley Evaluation & Assessment Research Center,
[11] S. Messick, Validity, In R.L. Linn (Ed.), Educational measurement (pp. 13-103), New York, American Council on Education, Macmillan, 1989
[12] R.J. Mislevy, M.R. Wilson, K. Ercikan, and N. Chudowsky, Psychometric Principles in Student Assessment, In D. Stufflebeam and T. Kellaghan (Eds.), International Handbook of Educational Evaluation, Dordrecht, The Netherlands, Kluwer Academic Press, 2001.
[13] R.J. Mislevy, L.S. Steinberg, and R.G. Almond, "On the structure of educational assessment" (to be published). Available http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/23/64/e6.pdf, 2003.
[14] NRC-National Research Council, Knowing what students know: The science and design of educational assessment, Committee on the Foundations of Assessment, 2001. Available: http://www.physics.ohio-state.edu/~jossem/REF/66.pdf
[15] N. Pellegrino, N. Chudowsky, and R. Glaser (Eds.), In NRC-National Research Council, Knowing what students know: The science and design of educational assessment, Committee on the Foundations of Assessment, 2001.
[16] K.K. Tatsuoka, "Rule Space: An approach for dealing with misconceptions based on item response theory", Journal of Educational Measurement, vol. 20, 1983, pp. 345-354.
[17] S. Toulmin, The uses of argument, Cambridge, England, University Press, 1958.
[18] D.M. Williamson, I.R. Katz, and I. Kirsch, An overview of the higher education ICT literacy assessment, 2004, ETS.
[19] B. Wright, M. Stone, Best test design, MESA Press, 1979.
[20] M. Wu, R. Adams, and M. Wilson, ACER ConQuest, Australia, Australia Council for Educational Research, 2005.