Computer-Based Assessment of Chinese Students' Component Skills of Problem Solving: A Pilot Study

Hao Wu and Gyöngyvér Molnár

Abstract—The objectives of the study is to explore the applicability of online assessment for detecting the component skills of problem solving and defining factors which impact the developmental level of student's problem solving skills in China. The sample of the pilot study was drawn from six grade students (age 11-13, N=50). The instruments of the study were conducted by four tests measuring problem solving, inductive reasoning, working memory and creativity, and a questionnaire, which focused on participants' demographic data, learning strategies, and ICT familiarity. The results and findings support the views that computer-based tests are applicable to measure students thinking skills and the component skills of problem solving in China at the age of 11-13, and proved the theoretical statement which is one's problem solving achievement is influenced by some specific thinking skills and background factors.

Index Terms—Computer-based assessment, problem solving skills, quantitative methods, thinking skills.

I. INTRODUCTION

Nowadays, our society, environment are keep changing all the times, the technologies of almost every industries are also developing quickly. This situation leads to people's content of applicable knowledge evolves rapidly. People are facing problems -small and large- almost everyday. The ability to solve problems timely and properly is gradually becoming one main factor for peoples' career and life. "Adapting, learning, daring to try out new things and always being ready to learn from mistakes are among the keys to resilience and success in an unpredictable world" [1]. Due to this situation, the research area problem solving has been hotly debated by the modern society and academic community.

The problem solving skills assessment can be divided into several different types. After years' development, interactive problem solving assessment has gradually became the mainstream. PISA 2012 creative problem solving and PISA 2015 collaborative problem solving, in essence, both belong to the interactive problem solving assessment scope. Interactive problem solving is "characterized by the interaction between a problem solver and the problem to generate and integrate information about the problem." [2] . In interactive problem solving assessment, relevant information needs to be actively generated, problem solvers

need to have direct interaction with the problem to uncover and discover relevant information [3]. It can be said that, the interaction between problems and problem solvers is the key part in the interactive problem solving assessment. But the traditional paper-and-pencil based assessment is not able to realized this kind of dynamic interaction, so it has gradually became obsolete in this area. Currently, the information and communication technologies provide new opportunities that can revolutionize the educational assessment and evaluation process [4]. Computer-based assessment is providing a unique assessment environment which the dynamic and interactive situations are available. And this kind of environment is cannot be provided by the use of paper-and-pencil instruments [5], [6]. Therefore, computer-based assessment has incontrovertibly replaced the position of paper-and-pencil assessment. It has became the most common assessment tool in the current problem solving assessment projects, which include PISA 2012 and PISA 2015 [1], [5], [7].

PISA 2012 problem solving assessment was focusing on the same area with our research, which is the individual interactive problem solving. Its result showed the development level of mainland Chinese (Shanghai) students' problem solving skills is in a high integral level. But there is one important problem about Chinese students' problem solving skills has been found. PISA 2012 had four subjects of assessment which are mathematics, reading, science and problem solving. The researchers of PISA have proved students' performance in problem solving is relevant with their performance in other three subjects [1]. Chinese students ranked as number 1 for all other three subjects. But for the problem solving assessment they are only ranked as number 6, which is far lower than their excepted level. And they have the almost largest difference between the real and the expected performance in the problem solving assessment. This result indicated Chinese students still have a great potential to improve their problem solving skills, and also proved the necessity for conduct a research project which focuses on Chinese students' problem solving skills.

II. LITERATURE REVIEW

A. Thinking Skills

The dynamic activities in the problem solving process have been described as information processing [8]. There are some researches have discussed or proved some thinking skills will be used in the information processing approach. Inductive reasoning, working memory, and creativity have been most frequently mentioned.

1) Inductive reasoning

Inductive reasoning has been proved as an important factor for the problem solving. Inductive reasoning skill is a general thinking skill, it is relevant with almost every higher-order cognitive skills and processes [9]-[11], and of course problem solving as well [1], [10], [12]. It will play an important role in most of the steps of the information processing approach such as the strategy selection and application, or some decision making activities.

2) Working memory

Working memory refers to a "brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks" [13]. Working memory has been shown has correlation between reasoning [13]-[15], intelligence [16], [17], and it also has impact on the problem solving progress. Problem solver needs to use working memory to store or transform some essential information during the information processing approach [18]. According to Sweller [19], working memory can be used to store the information like 1) the strategies which potentially be used; 2) the known variables such as the current and previous strategies the solver selected and the states which have occurred; and 3) the unknown variables which are needed in the information processing approach. Of course, the information in working memory can be transformed (e.g. from unknown to known), added, or deleted [19]. The features like capacity or operability of one's working memory are influencing his/her performance in cognitive processes which include problem solving [20]. Passolunghi & Siegel [21] reported that the working memory deficit will cause difficulties during the information processing approach and may leads to failure for problem solving.

3) Creativity

Creativity means the ability for take "the process of having original ideas that have value" [22], it is a one of the most important 21st century skills [23], which means it has important meaning for people's study, work and life in our new century [24]. Creativity is also a thinking skill which has been reported has correlation with the problem solving process [23], [25]. Creative ideas are often required in the problem solving process [24]. For example, this skill can help the problem solver when they are creating or selecting their strategies. In addition, based on students' performance in PISA 2012, the researchers found that lack of creativity may lead to the problem solver less successful when they are dealing with some problems they are not familiar with [1].

Based on the literature review, a preliminary hypothesis for students' problem solving skills structure has been built. In this stage, we assume inductive reasoning, working memory and creativity these three thinking skills will influence one's problem solving achievement; and we assume there are internal relationships between these three thinking skills.

B. Non-cognitive Variables

Problem solving also can be impacted by some non-cognitive factors, the following parts will illustrate some typical cases.

Motivation: Motivation has already been proved has important influence in people's mental or practical activities [26], [27]. Motivation can "influence the way people attempt to understand and control a dynamic system" [8], and of course problem solving will also be impacted by solver's motivation according to previous description. According to Frensch & Funke's [8] experiment result, the students who have higher motivation showed better understanding for a system and higher efficiency in the problem solving process.
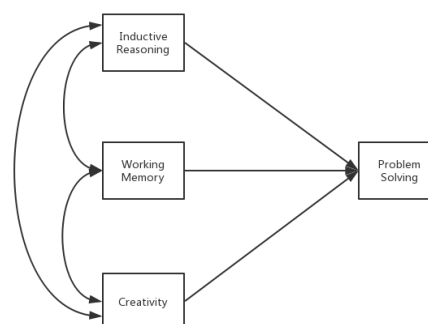

Fig. 1. Preliminary hypothesis for problem solving component skills.

Age & Gender: The ability for students' problem solving is developing with their age [10]. According to Molnár, Greiff and Csapós [10] work, grade 5 to grade 8 is the most important time stage for students' problem solving skill development. In addition, in most of countries, students with different gender has different performance in the problem solving assessment. According to the data from PISA 2012, generally boys showed better performance than girls in the problem solving assessment [1].

ICT literacy: ICT literacy also has been considered as one of the 21st centenary skills [23]. With the development of our world, there are increasing number of problems which occurred in people's daily life are relevant with the electric devices. People with high ICT literacy will have high probability to successfully solve this kind of problems [23]. Besides, according to the previous part, the interactive problem solving assessment only can be realized by computers. Therefore, students' performance in a problem solving assessment will inevitably impacted by their ICT literacy [1], [5].

III. AIMS

The aim of this study is to detecting and finding the factors impacting the developmental level of problem solving, and building model to present the system of constitution for Chinese students' problem solving component skills. But in the start stage of this study, there are some issues have to be confirmed at first. Firstly, online thinking skills assessment is not very common in China, therefore we need to confirm online assessment is applicable in China, and make sure our online assessment platform can be run in the Chinese network environment. Secondly, whether the thinking skills assessment tasks which going to be used in this study are reliable in the Chinese culture is unknown at the first stage. Thus measure the reliability of each test is an essential work at this stage. Last but not least, the pilot test also provides an opportunity for doing a preliminary testing of the theoretical study results.

To sum up, the aims for doing this pilot can be summarized as following:

1) to explore the applicability of online assessments in China,
2) to test the reliability of every cognitive tests involved in the project (problem solving, inductive reasoning, working memory, creativity)
3) to achieve a preliminary conclusion to verify the theoretical study and make guidance for further study.

IV. METHODS

A. Participants

The were 50 Chinese students (27 boys; 23 girls) have attended the pilot test. All the participants are six graders (age mean=12.28, standard deviation (SD=.50). According to the Chinese educational system, the participants were in the last year of their primary school.

B. Instruments

The pilot test consisted by four contests which focusing on students thinking skills (1. problem solving, 2. inductive reasoning, 3. working memory and 4. creativity) and one background questionnaire.

The problem solving assessment instrument consisted by 12 items which adopted from a mature problem solving assessment system named as MicroDYN [28]. The first stage of the assessment was an introduction section which include text and video based instructions and a trial task. The assessment contains two phrases. In the first phrase, students need to interact with the system, and find the interconnection between each variables through the interaction (this phrase named as knowledge acquisition; [28]). The second phrase of assessment requires students to apply their knowledge to control the system to a required state (this phrase named as knowledge application; [28]). Students had three minutes to make their answers for each item. In addition, some test-taking motivation questions have been added into the problem solving assessment progress to monitor students' motivation during this assessment. The motivation questions asked students whether they think this assessment is interesting, and entertaining; and whether they enjoy this assessment. The questions have 7 scales from 1- strongly disagree to 7- strongly agree.

Fig. 2. Sample items for the inductive reasoning test.

The inductive reasoning test administered in China consisted of 53 items. The test items were adopted from Pásztor, *et al.*'s [29] work. We used pictures and numbers as stimuli. Students need to detect, find and apply the certain rules between the given pictures or numbers to make their answers.
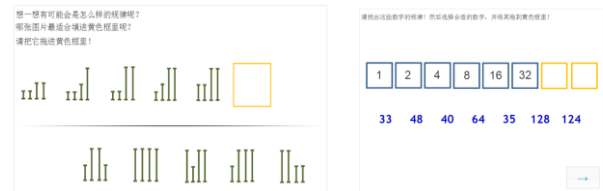

Fig. 3. Sample items for the inductive reasoning test.

The working memory test contains 11 items. In this test, we showed some certain pictures which constituted by black and white squares. Each picture lasts for three seconds on the screen. After that, students need to represent the pictures in the answer zone.

Fig. 4. Sample item for the working memory test.

The creativity test was adopted from Pasztor, Molnár & Csapós [24] work. Students need to provide every single possible usage they can image for three daily necessities (match, cup and toothbrush). There were also three picture meaning tasks have been included, students need to write every meaning they can image from three different pictures into the textboxes. Students had three minutes to make their answers for each item.

Fig. 5. Sample item for the creativity test.

Fig. 6. Stimuli for picture meaning tasks.

The background questionnaire in this pilot test was mainly focus on students demographic data (e.g. gender, age, etc.), daily ICT usage situation and regularly applied learning strategies. Part of the questions was adopted from the existing questionnaires from PISA [1], [7], [30].

C. Procedures

The test was carried out by the eDia (Electronic Diagnostic Assessment; [4]) platform in the school's ICT room. Test completion was divided into three sessions, each lasting approximately 45 minutes. In session 1, students worked on the inductive reasoning test. In session 2 students had to complete the problem solving test, in session 3 the working memory test, the creativity test and the questionnaire. All the items in the pilot test were adopted into simplified Chinese.

D. Data Analysis

Structural equation modeling (SEM) was used to analyze the dimensionality of problem solving and the relationship between each thinking skills. All the models were computed by Mplus version 5 [31]. CFI (Comparative Fit Index), TLI (Tucker-Lewis Index), SRMR (Standardized Root Mean Square Residual) and RMSEA (Root Mean Square Error of Approximation) have been calculated by Mplus to indicate the model fit. In addition, independent t-test has been used to analyze the relationship between thinking skills and background factors. The t-test analyses were carried out by SPSS version 22.

V. RESULTS

A. Psychometric Properties and Reliability Testification

The pilot test has been successfully implemented. According to the interviews after the assessment, students had no problem with the operation of our computer-based assessments. The mean values, standard deviations and reliability (Cronbach's alpha) for each subtest were as following:

TABLE I: MEAN, STANDARD DEVIATION AND CRONBACH'S ALPHA FOR EACH COGNITIVE ASSESSMENTS

Subtest	M	SD	Cronbach's alpha
Problem Solving	47.73%	21.78%	0.72
Inductive reasoning	80.33%	8.63%	0.75
Working Memory	26.9%	25.1%	0.85
Creativity	2.91	4.02	0.90

Beside the results from the cognitive assessments, the motivation questions in the problem solving assessment also provided information which regarding student's test-taking motivation. We assumed the highest motivation (choose 7-strongly agree for every single question) as 100%. The average response for the students is 83.85% (SD=16.97%). The results demonstrated students' high motivation in the problem solving assessment.

The primary aim of our pilot is to explore the applicability of online assessments in China and to test the reliability of every cognitive tests involved in the project. As indicated above, the problem solving and inductive reasoning tests showed acceptable internal consistencies; in the meanwhile, the internal consistencies for working memory and creativity tests were satisfactory. Our results proved computer-based assessment is feasible and reliable in China, and our cognitive tests are also reliable to measure Chinese students' thinking skills.

B. Dimensionality of Problem Solving

The problem solving assessment contains two phrases, knowledge acquisition and knowledge application. An one-dimensional and a two-dimensional measurement model have been built to confirm the dimensionality of problem solving assessment result. Both of the one-dimensional and two-dimensional model showed good model fits.

In the one-dimensional model, two sub-factors (knowledge acquisition and knowledge application) have been considered as one general factor problem solving; while the two-dimensional model creates a latent variable as problem solving which correlated with knowledge acquisition and knowledge application these two factors. Based on the above table, both of these two models got good model fits. In addition, the chi-square test for difference testing computed by Mplus showed there is no significant difference between these two models (p>.05). The analysis results showed in our case, problem solving can be explained as both one-dimensional and two-dimensional model. In this paper, two-dimensional model will be selected for doing analysis because it has sightly better model fits and it is able to show more details.

TABLE II: GOODNESS OF FIT INDICES FOR TESTING DIMENSIONALITY OF PROBLEM SOLVING

Model	Chi-square	df	RMSEA	CFI	TLI
1-dimensional	15.257	16	0.000	1.000	1.011
2-dimensional	14.280	15	0.000	1.000	1.012

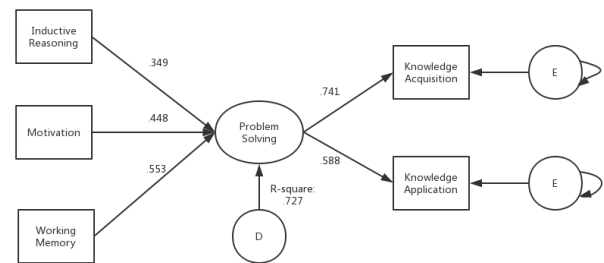### C. SEM Analysis of Component Skills of Problem Solving

Fig. 7. Structural model of problem solving (two-dimensional).

The SEM model with two-dimensional problem solving got good model fit (RMSEA=0.000; SRMR=0.025; CFI=1.000; TLI=1.110). The model showed problem solving can be explained by knowledge acquisition and application these two factors (r=.588-.741; p<.001). Inductive reasoning, test-taking motivation and working memory showed significant positive influence (r=.349-.553, p<.05) on one's problem solving achievement in this model. The influence from working memory turned out to be the strongest. The results effectively proved these three factors' importance during the problem solving approach. But creativity did not show significant correlation with problem solving, and significant correlation between inductive reasoning and working memory can not be found, which *partly denied* our preliminary hypotheses.

D. Relationships between Thinking Skills and Background Factors

There was no significant gender difference can be found. Girls performed better than boys in creativity (t=1.072, p>.05), working memory (t=.362, p>.05) and problem solving (t=.103, p>.05), but lower than boys in inductive reasoning (t= -.886, p>.05). But none of these differences were statistically significant, which means basically boys and girls performed almost equal in our pilot test. Moreover, we found students' ICT usage situation (include how many electric devices they have in their home and how long time they use electric devices in school and home) showed moderate but significant correlation with their performance in problem solving assessment (r=.384, p<.05). Which

indicated students' ICT literacy was influencing our assessment in a certain but not strong level.

Another important finding we have got is students' regularly applied learning strategies are influencing their think skills development. In the background questionnaire we have listed 13 learning strategies which can be grouped into three categories which are 1) elaboration strategies (strategies for link their learning content with their previous knowledge or real-life), 2) memorisation strategies (strategies for remember everything without thinking) and 3. control strategies (strategies for doing self-control or time-control). These 13 learning strategies were referenced from PISA 2003 [30] questionnaires with some modifications. Students need to choose the frequency (5-scale: Never, Rarely, Sometimes, Often, Always) for using the listed learning strategies in their daily study. Based on students' responses, they have been divided into different groups by their most commonly used learning strategies. Significant differences have been found between different groups' students' cognitive tests performance. The results indicated students who preferred memorization strategies in their study had significant lower performance in the inductive reasoning test than who did not prefer memorization (t= -2.942, $p<.05$). In the meanwhile, students who preferred control strategies had shown significant better performance in problem solving ($t=2.194$, $p<.05$), inductive reasoning ($t=2.806$, $p<.05$) and creativity ($t=2.197$, $p<.05$) than their peers. The results proved regularly applied learning strategies' importance in students' thinking skills development. Results also showed the possibility that instructors might can help students to develop think skills by train and encourage students to apply more effective learning strategies.

VI. DISCUSSIONS

The aims for conducting this pilot have been illustrated in section 3. The reliability of each tests turn out to be high. Which indicated computer-based assessment, and our online assessment platform (eDia) is applicable in China; and the tests of problem solving, inductive reasoning, visual memory and creativity are reliable to measure students' thinking skills in China. The result gave a positive answer to the basic but most important question, which is whether our study is possible to be conducted in the China context. Furthermore, the results proved the theoretical statement: one's problem solving achievement is influenced by some specific thinking skills and background factors. In our case, we preliminary confirmed the influence from some think skills such as inductive reasoning and working memory and some non-cognitive factors such as ICT usage and learning strategies to students' problem solving achievement. Findings proved the possibility for building model to present the cognitive system of Chinese students' problem solving skills. To conclude, the findings effectively testified the feasibility of this project, and built solid foundation for future study.

VII. LIMITATIONS AND FUTURE WORK

There are some limitations need to be noted. Firstly, analysis showed some tests' difficulty level were not perfectly suitable for assess Chinese students' thinking skills in this age group (e.g. Rasch model analysis showed some items in the inductive reasoning test were too easy for the participating students, meanwhile, some items in the problem solving test were a bit difficult for them). Moreover, the small sample size might impact our data analysis accuracy. In addition, in this assessment, all the participants were from same grade, which caused troubles for analyze students' think skills development in different age and grade. A large-scale assessment (500 participants according to the plan) is going to be conducted soon. The limitations which mentioned above will be noted and addressed.

REFERENCES

[1] OECD, "PISA 2012 Results: Creative problem solving−Students' skills in tackling real-life problems," volume V, OECD, Paris, France, 2014.

[2] S. Greiff, D. Holt, and J. Funke, "Perspectives on problem solving in cognitive research and educational assessment: analytical, interactive, and collaborative problem solving," *Journal of Problem Solving (The)*, vol. 5, pp. 71-91, 2013.

[3] OECD, "PISA 2012 field trial problem solving framework," OECD, Paris, France, 2010.

[4] B. Csapó, A. Lőrincz, and G. Molnár, "Innovative assessment technologies in educational games designed for young students," *Assessment in Game-Based Learning: Foundations, Innovations, and Perspectives*, Springer, New York, pp. 235-254, 2012.

[5] S. Greiff, A. Kretzschmar, J. Müller, B. Spinath, and R. Martin, "Computer-based assessment of Complex Problem Solving in educational contexts and how it is influenced by students' level of Information and Communication Technology literacy," *Journal of Educational Psychology*, 2014.

[6] G. Molnár, S. Greiff, S. Wüstenberg, and A. Fischer, "Empirical study of computer-based assessment of domain-general complex problem-solving skills," *Educational Research and Innovation*, pp. 125-140, 2017

[7] OECD, "PISA 2012 results: What students know and can do, student performance in mathematics, reading and science (Volume I)," OECD, Paris, France, 2014.

[8] J. Funke and P. A. Frensch, "Complex Problem Solving: The European Perspective," 1995.

[9] B. Csapó, "The development of inductive reasoning: Cross-sectional assessments in an educational context," *International Journal of Behavioral Development*, vol. 20, no. 4, pp. 609-626, 1997.

[10] G. Molnár, S. Greiff, and B. Csapó, "Inductive reasoning, domain specific and complex problem solving: Relations and development," *Thinking Skills and Creativity*, vol. 9, pp. 35-45, 2013.

[11] T. Schubert, D. E. Astle, and M. J. Molen, "Computerized training of non-verbal reasoning and working memory in children with intellectual disability," *Training-Induced Cognitive and Neural Plasticity*, 2012.

[12] W. Tomic, "Training in inductive reasoning and problem solving," *Contemporary Educational Psychology*, vol. 20, no. 4, pp. 483-490, 1995.

[13] A. Baddeley, "Working memory," *Science*, vol. 255, no. 5044, pp. 556-559, 1992.

[14] S. Söderqvist, S. B. Nutley, J. Ottersen, K. M. Grill, and T. Klingberg, "Computerized training of non-verbal reasoning and working memory in children with intellectual disability," *Frontiers in Human Neuroscience*, vol. 6, 2012.

[15] S. Greiff, A. Fischer, M. Stadler, and S. Wüstenberg, "Assessing complex problem-solving skills with multiple complex systems," *Thinking & Reasoning*, pp. 1-27, 2015

[16] P. L. Ackerman, M. E. Beier, and M. O. Boyle, "Working memory and intelligence: The same or different constructs?" 2015.

[17] S. M. Jaeggi, M. Buschkuehl, J. Jonides, and W. J. Perrig, "Improving fluid intelligence with training on working memory," in *Proc. the National Academy of Sciences,* vol. 105, no. 19, pp. 6829-6833, 2008.

[18] K. Oberauer, H. M. Süß, R. Schulze, O. Wilhelm, and W. W. Wittmann, "Working memory capacity—facets of a cognitive ability construct," *Personality and Individual Differences,* vol. 29, no. 6, pp. 1017-1045, 2000.

[19] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognitive science*, vol. 12, no. 2, pp. 257-285, 1988.

[20] K. Fukuda and E. K. Vogel, "Human variation in overriding attentional capture," *The Journal of Neuroscience*, vol. 29, no. 27, pp. 8726-8733, 2009.

[21] M. C. Passolunghi and L. S. Siegel, "Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving," *Journal of Experimental Child Psychology,* vol. 80, no. 1, pp. 44-57, 2001.

[22] K. Robinson, *Out of Our Minds: Learning to Be Creative*, Chichester: Capstone, 2009.

[23] M. Binkley, O. Erstad, J. Herman, S. Raizen, M. Ripley, M. Miller-Ricci, and M. Rumble, "Defining twenty-first century skills," *Assessment and Teaching of 21st Century Skills,* pp. 17-66, Springer Netherlands, 2012.

[24] A. Pásztor, G. Molnár, and B. Csapó, "Technology-based assessment of creativity in educational context: the case of divergent thinking and its relation to mathematical achievement," *Thinking Skills and Creativity,* vol. 18, pp. 32-42, 2015

[25] N. Herrmann "Creative problem solving," *IEEE Potentials Magazine*, pp. 4-9, 1995.

[26] V. H. Vroom, *Work and Motivation. 1964*, NY: John Wiley &sons, vol. 45, 1964.

[27] J. R. Hackman and G. R. Oldham, "Motivation through the design of work: Test of a theory," *Organizational Behavior and Human Performance,* vol. 16, no. 2, pp. 250-279, 1976.

[28] S. Greiff, A. Fischer, S. Wüstenberg, P. Sonnleitner, M. Brunner, and R. Martin, "A multitrait−multimethod study of assessment instruments for complex problem solving," *Intelligence,* vol. 41, no. 5, pp. 579-596, 2013.

[29] A. Pásztor, G. Molnár, E. Korom, B. Németh, and B. Csapó, "Online assessment of inductive reasoning and its predictive power on inquiry skills in science," presented at 17th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI), Tampere, Finland, August 29-September 2, 2017.

[30] OECD, "Problem solving for tomorrow's world: First measures of cross-curricular competencies from PISA 2003," OECD Publishing, 2004.

[31] L. K. Muthén, and B.O. Muthén, *Mplus User's Guide,* Fifth Edition, Los Angeles, CA: Muthén & Muthén, 2007.

Hao Wu received the B.S degree in mathematics from Anhui University, Hefei, China in 2012, and the master degree (with Honours) in computer science from the Australian National University, Canberra, Australia in 2014. He is currently a PhD student in the Doctoral School of Education, University of Szeged, Hungary. His main research areas are ICT in education, computer-based assessment and problem solving skills.

Gyöngyvér Molnár is a professor of education at the University Szeged, Hungary. Her main research areas are technology-based assessment, educational measurement, especially issues of item response theory and problem solving.