Diagnostic Assessment for Improving Teaching Practice

Yuan Sun and Masayuki Suzuki

Abstract—**Most of current tests only report students' total test scores, or T-scores, which are almost no use for providing diagnostic information to inform teachers of their students' strengths and weaknesses. Additionally, it has been shown that most teachers have difficulty in using assessment to improve their teaching. Recently, cognitive diagnostic assessment (CDA) which is designed to measure specific knowledge structures and processing skills in students has attracted a great deal of attentions. In this paper, we apply a CDA approach in fraction problems to 144 sixth grade students in an elementary school in Japan. We show how CDA can provide detailed information about students' strengths and weaknesses and discuss the applicability of CDA for providing effective feedback for teachers to improve their teaching practice.**

Index Terms—**Cognitive diagnostic assessment, attribute mastery probability, effective feedback, fraction problems.**

I. INTRODUCTION

Assessment provides teachers with information about such things as whether students have learned enough or what to teach them. Teachers are expected to use assessment results to improve their teaching and organize a sound instruction plan for facilitating students' educational development [1].

Despite the importance of assessment in education today, few teachers receive much formal systematic training in assessment design or analysis [2]. As a result, teachers rely heavily on the assessment offered by the publisher of their textbooks or instructional material, or their own constructed classroom tests which are primarily used for assigning students' grades. It has been pointed out that teachers' assessment literacy is low and most teachers have difficulty in using assessment appropriately [3], [4]. To use assessment to improve instruction and student learning, teachers need more training [5], [6]. Needless to say, it is important to plan and deliver training programs for enhancing teachers' level of assessment competencies. On the other hand, to develop suitable assessment tool to support teachers for improving their teaching practice should be effective and important as well.

In this study we aim to introduce cognitive diagnostic assessment and its method for providing feedback effectively and show the assessment has the potential to provide useful information for teachers to improve their teaching. Feedback is conceptualized here as information provided by an assessment regarding aspects of students' performance or understanding.

II. COGNITIVE DIAGNOSTIC ASSESSMENT IN FRACTION PROBLEMS

A. Cognitive Diagnostic Modeling

Most of current tests, which only report a small number of content-based subscores, total scores, or T-scores, are almost no use for providing diagnostic information to inform teachers of their students' strengths and weaknesses. In recent years, cognitive diagnostic assessment (CDA) which is designed to measure specific knowledge structures and processing skills in students has been particularly attracting a great deal of attention [7], [8]. CDA combines theories of cognition of interest with statistical models intended to make inferences about students' mastery of so called "attributes", which refer to students' "knowledge", "cognitive processes", "skills" or "strategies" in a particular domain [9]. Through fine-grained diagnostic reporting of students' attribute mastery profiles, CDA provides more detailed information concerning whether or not, or to what extent students have mastered each of a group of specific defined attributes, rather than assigning each student only one single score. These attributes are exactly what affect students' learning performance and their understanding and are exactly what teachers need to help students to master.

In a cognitive diagnostic model, suppose there are k attributes in a particular domain, each student's mastery profile is formalized as a vector $\alpha = (\alpha_1, \ldots, \alpha_k)$; α_k indicates the student's true ability along attribute k: $\alpha_k = 1$ indicates mastery of attribute k and $\alpha_k = 0$ indicates nonmastery of this attribute. Only certain attributes are measured by each item; information relating items to attributes is typically given by the Q-matrix (item j ×attribute k) indicating which skills and knowledge are involved in solving each item [9]. Letting $j = 1, \ldots, J$ index the items in a test, the [j, k] entry of the Q-matrix (denoted q_{jk}) is equal to one if item j measures attribute k; and zero otherwise.

Attributes necessary for the domain need to be specified first by experts in that area, by specifying what the key knowledge or skills are for students to understand the content in their learning process. After that, test items and Q-matrix need to be constructed. Once the Q-matrix has been specified and items have been administered in a test, the items are calibrated to one of cognitive diagnostic models. In this study a DINA (deterministic input, noisy ''and'' gate) model [10], [11] was applied in solving fraction problems.

B. A Fraction Diagnostic Test

We have developed eight attributes required in solving fraction problem according to the "Japanese government curriculum guidelines for teaching [12]" (Table I) (see [13], [14] for details). The numbers in the parentheses in the table denote grades where students are supposed to study the skills,

according to the guidelines.

TABLE I: EIGHT ATTRIBUTES DEFINED IN SOLVING FRACTION PROBLEMS

Attribute	descriptions
A1	understanding the meaning of fraction (3rd grade)
A2	changing fraction forms(4th)
A3	reducing to a common denominator (5th)
A4	reducing to the lowest denominator (5th)
A5	adding and subtracting fractions with a common denominator (3rd)
A6	multiplying (6th)
A7	dividing (6th)
A8	deriving fractions from statements (3rd)

We comprised a diagnostic test of 35 items and administered it to 144 sixth grade students in an elementary school in Tokyo.

III. RESULTS AND DISCUSSION

A. Students' Individual Attribute Mastery Profiles

Table II shows examples of five students' attribute mastery probabilities estimated based on the DINA model who have the same test scores of 22 (number of correct answers). We can see the student 2 masters A5 fully but does not master A4 at all. On the other hand, the student 3 masters A4 fully but does not master A5.

Although the five students have exactly the same test score, we can see they have different achievements in each attribute. The radar charts for their attribute mastery probabilities, which gives us a visual presentation more easily understood, are shown in Fig. 1. It was clear that students even with exactly the same test scores could have totally different attribute mastery profiles. These data cannot be obtained from the current typical educational tests.

Each individual student's attribute mastery probabilities were reported which should be useful information for teachers to know about each student's detailed knowledge state and give them appropriate guidance individually for their future remedial work. For example, for student 2, who does not master the skills of "changing fractions to equivalent fractions with a common denominator" (A3) and "reducing a fraction" (A4), it might be effective to take them back to go through again the meanings of the concept "reduction to common denominator" and "reduction to the lowest fraction", and support them to learn about "the greatest common divisor" and "the least common multiple".

Teachers do not obtain this kind of diagnostic information for individual students from the current typical educational tests which are more focusing on students' total scores. CDA can give teachers detailed information and support them to organize an appropriate instruction plan for each student.

B. Group Attribute Mastery Profiles

Based on students' knowledge states, we also performed a hierarchical cluster analysis using Ward's method to categorize students for teachers to know the whole situation and adjust their teaching practice. As a result, 144 students were classified into four groups with different types of attribute mastery profiles. Table III shows the four clusters' profiles, the sizes and mean scores of each cluster.

TABLE II: EXAMPLES OF FIVE STUDENTS' ATTRIBUTE MASTERY PROBABILITIES WHO HAVE SAME NUMBER OF CORRECT ANSWERS

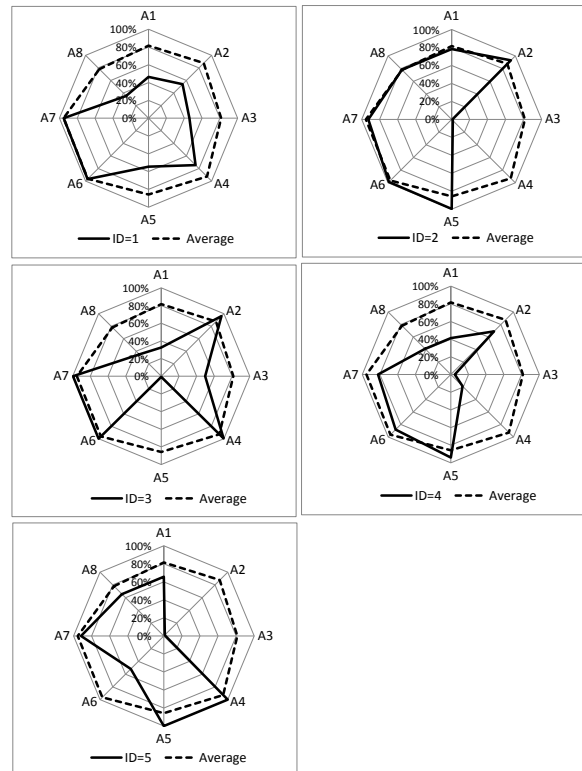| ID	A1	A2	A3	A4	A5	A6	A7	A8
1	0.47	0.54	0.46	0.75	0.54	0.97	0.95	0.35
2	0.78	0.93	0.01	0.02	1.00	0.99	0.93	0.78
3	0.33	0.96	0.50	1.00	0.01	1.00	1.00	0.35
4	0.41	0.69	0.05	0.19	0.94	0.88	0.82	0.41
5	0.66	0.02	0.02	1.00	1.00	0.52	0.92	0.66

Fig. 1. Radar charts of the five students' attribute mastery probabilities.

TABLE III: FOUR CLUSTERS WITH DIFFERENT TYPES OF ATTRIBUTE MASTERY PROBABILITIES

Clusters	Mean scores	Attribute mastery probabilities							
		A1	A2	A3	A4	A5	A6	A7	A8
Cluster 1 (n = 96)	30.9	0.96	0.97	0.98	1.00	0.99	0.98	0.99	0.99
Cluster 2 (n = 12)	27.8	0.52	0.99	1.00	1.00	1.00	1.00	1.00	0.04
Cluster 3 (n = 14)	17.9	0.62	0.41	0.04	0.39	0.96	0.77	0.65	0.63
Cluster 4 (n = 22)	22.7	0.48	0.68	0.47	0.94	0.15	0.98	0.98	0.64

Fig. 2 gives a more visual display of the four types of overall mastery status for teachers to have a clear grasp and understanding about the class overall, which could be useful for teachers to adjust their teaching contents and activities in group guidance.

It is clear that students in the first cluster have mastered all attributes completely, which indicates they have already acquired knowledge or skills necessary for solving fraction

problem. It can be also seen that about 67% of all the students in the school in our study have achieved the level required by "Japanese government curriculum guidelines" (Table III). While students in the second cluster haven't mastered A8 at all, but all other attributes almost perfectly. It indicates that these students might have problems with reading comprehension or lack skills to derive expression from word problems rather than fraction calculation itself. We also see that students in the third cluster have their mastery probabilities of almost all the attributes except only for A5 below 80%. In other words, it is indicated that about 10% students haven't mastered skills required in solving fraction problem at all and more comprehensive supplementary instructions for these students need to be provided.
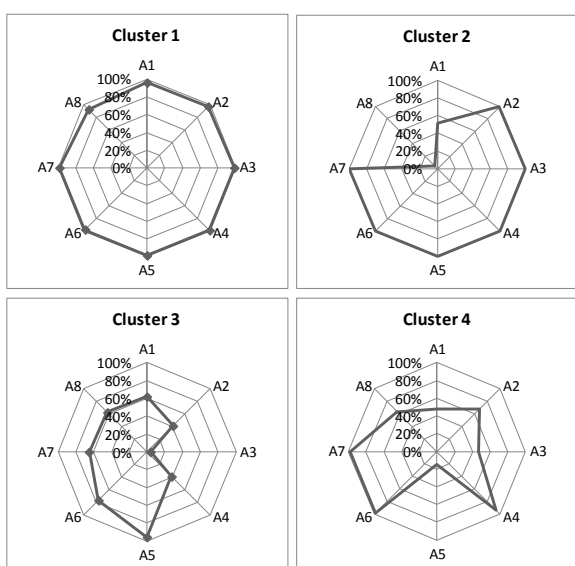

Fig. 2. Radar charts of the four clusters' attribute mastery profiles.

Especially students in the fourth cluster is worthy of notice. They are capable of multiplying and dividing fractions which are the skills required in 6th grade. Nevertheless, they are not capable of adding and subtracting fractions with a common denominator (A5) which are the easier skills supposed to be acquired in 3rd grade. In fact, even students in the 3rd cluster, which has the lowest average score among the four clusters, have mastered the skill A5 almost perfectly. These students in the fourth cluster might have conducted addition and subtraction of fractions under a wrong rule: for example applying rules of multiplying or dividing for adding or subtracting by mistake (e.g., $2/5 + 1/5 \rightarrow 2/25$). Teachers need to realize the problems for these students and amend their wrong knowledge of calculation algorithms adequately.

As shown above, categorization of students based on their knowledge states makes it clearer and easier to understand the situation about the whole class. These concise summaries of the results might be more useful sometimes than the information of each individual student's profile for teachers, especially for those teachers who are teaching many students and difficult to know and deal with student's problem individually.

IV. CONCLUSION

How to assist teachers to improve their teaching practice is a very important issue in education. In this study we introduced a cognitive diagnostic fraction test and CDA method for providing feedback information to teachers in an elementary school. After showing the results to the three homeroom teachers, we also conducted a questionnaire survey to ask about the effectiveness and usefulness of CDA's feedback results. We received positive answers from all the teachers considering diagnostic information of students' attribute mastery probabilities, which are not possible to get from the current typical educational tests, as effective feedback information. How assessment could aid or support teaching activities is a challenging but crucial issue nowadays. Our study shows CDA could be a useful approach. To use CDA in practice, diagnostic score reporting strategies for effective feedback need to be developed [15]. In fact, attribute mastery probabilities is incomprehensible because the meaning of probability is difficult to understand. Hence, feedback should be provided in the form of written documents containing statement that describe assessment standard and students' achievement. How to effectively communicate such complex and detailed information on educational tests, including feedback of the reasons behind failures and history of misunderstanding in learning process, and provide guidelines for teaching improvement is a very difficult task, and further research is required.

REFERENCES

[1] American Federation of Teachers, National Council on Measurement in Education, National Education Association (AFT, NCME, NEA), "Standards for teacher competence in the educational assessment of students," *Educational Measurement: Issues and Practice*, vol. 9, pp. 30-32, 1990.

[2] W. D. Schafer, "Assessment literacy for teachers," *Theory into Practice*, vol. 32, pp. 118-126, 1993.

[3] C. DeLuca and D. Klinger, "Assessment literacy development: Identifying gaps in teacher candidates' learning," *Assessment in Education: Principles, Policy and Practice*, vol. 17, pp. 419-438, 2010.

[4] B. S. Plake and J. C. Impara, "Teacher assessment literacy: What do teachers know about assessment?" in *Handbook of classroom assessment: Learning, achievement, and adjustment*, G. D. Phye Ed., Academic Press, 1997, pp. 53-68.

[5] J. H. McMillan, S. Myran, and D. Workman, "Elementary teachers' classroom assessment and grading practices," *Journal of Educational Research*, vol. 95, pp. 203-213, 2002.

[6] W. T. Rogers, "Educational assessment in Canada: Evolution or extinction?" *Alberta Journal of Educational Research*, vol. 37, pp. 179-192, 1991.

[7] Y. Lee and Y. Sawaki, "Cognitive diagnosis approaches to language assessment: An overview," *Language Assessment Quarterly*, vol. 6, pp. 172-189, 2009.

[8] A. A. Rupp, J. Templin, and R. A. Henson, *Diagnostic measurement: Theory, methods, and application*, New York: Guilford, 2010.

[9] K. Tatsuoka, "A probabilistic model for diagnosing misconceptions in the pattern classification approach," *Journal of Educational Statistics*, vol. 12, pp. 55-73, 1985.

[10] J. de la Torre, "DINA model and parameter estimation: A didactic," *Journal of Educational and Behavioral Statistics*, vol. 34, pp. 115-130, 2009.

[11] B. W. Junker and K. Sijtsma, "Cognitive assessment models with few assumptions, and connections with nonparametric item response theory," *Applied Psychological Measurement*, vol. 25, pp. 258-272, 2001.

[12] Ministry of Education, Culture, Sports, Science and Technology. (2008). Primary school curriculum guidelines. [Online]. Available:

http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2009/04/21/1261037_4.pdf

[13] Y. Sun, M. Suzuki, and S. Kakinuma, "Effective feedback for self-regulated learning: Applying cognitive diagnostic assessment," *Advances in Education Research*, vol. 7, pp. 140-145, 2012.

[14] T. Takahashi, Y. Sun, and S. Kakinuma, "Development of attributes of cognitive diagnostic test in fraction calculations," in *Proc. 2011 conference of the Japan Association for Research on Testing*, 2011, pp. 220-221.

[15] M. Roberts and M. Gierl, "Developing score reports for cognitive diagnostic assessments," *Educational Measurement: Issues and Practice*, vol. 29, pp. 25-38, 2010.

Yuan Sun is an associate professor at the Information and Society Research division, National Institute of Informatics. Her current research interests include measurement in education and psychology, especially test theory, learning analytics, and institutional multi-dimensional evaluation.
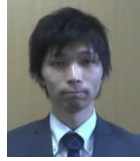
Masayuki Suzuki is a project researcher at the Information and Society Research division, National Institute of Informatics. He has a Ph.D. in Education from Tokyo University. His major field of study is educational psychology, and his research interests are educational evaluation and assessment.