Wong, L.-H. et al. (Eds.) (2013). Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education

# Designing Effective Feedback for Cognitive Diagnostic Assessment in Web-based Learning Environment

**Yuan SUN*, Masayuki SUZUKI & Tetsuya TOYOTA**
*National Institute of Informatics, Japan*
*yuan@nii.ac.jp

**Abstract:** Assessment is useful for students to improve their learning and for teachers to adjust their teaching practice. However, most traditional assessments do not provide useful information to improve learning and 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 learners' strengths and weaknesses and discuss the applicability of web-based CDA for providing effective feedback.

**Keywords:** cognitive diagnostic assessment, web-based assessment, feedback

## 1. Introduction

Assessment is helpful to inform students about how their present state relates to goals and standards and what are the learning tasks to work on (e.g., Clark, 2012). Assessment can also provide teachers with information about how students have learned and what they need to teach them. Teachers can use assessment results to organize a sound instruction plan for facilitating students' educational development or evaluate instruction to improve their teaching skills (AFT, NCME, NEA, 1990).

In current school education, to assess students' present state of learning related to learning goals, teachers rely heavily on their own constructed classroom tests which are primarily used for assigning students' grades. However, there are a number of problems with this assessment when applied to formative assessment. Firstly, students have difficulty in understanding feedback message (Higgins, Hartley, & Skelton, 2001). When teachers transmit feedback information to students those messages have to be easily decoded and translated into action. Secondly, few teachers receive much formal systematic training in assessment design or analysis (McMillan, Myran, & Workman, 2002). It has been pointed out that teachers' assessment literacy is low and most teachers have difficulty in using assessment appropriately (DeLuca & Klinger, 2010). Thirdly, traditional assessment procedures lack detailed diagnostic information about students' academic achievement and do not provide meaningful information to guide remedial education (Lee & Sawaki, 2009).

In this paper, we aim to introduce cognitive diagnostic assessment (CDA) and its method for providing feedback effectively and show the assessment has the potential to provide useful information for students to improve their learning and for teachers to adjust their teaching contents and activities, especially when assessment is implemented in web-based learning environment.

## 2. Cognitive Diagnostic Assessment in Fraction Problems

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 about students' strengths and weaknesses. In recent years, CDA which is designed to measure specific knowledge structures and processing skills in students has been particularly attracting a great deal of attention. 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 (Tatsuoka, 1985). 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.

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. We have developed eight attributes required in solving fraction problem according to the "Japanese government curriculum guidelines for teaching" (Table 1) (for details, see Takahashi, Sun, & Kakinuma, 2011). The numbers in the parentheses in the table denote grades where students are supposed to study the skills, according to the guidelines. We comprised a diagnostic test of 35 items and administered it to 144 sixth grade students in an elementary school in Tokyo. Table 2 shows examples of five students' attribute mastery probabilities who have the same number of correct answers. In this paper, DINA (deterministic input, noisy "and" gate) model (Junker & Sijtsma, 2001) was applied to estimate attribute mastery probabilities. 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. 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.

Table 1: Eight attributes defined in solving fraction problems.

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

Table 2: Examples of five students' attribute mastery probabilities who have the same score.

| Student | Total score | A1 | A2 | A3 | A4 | A5 | A6 | A7 | A8 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 22 | 0.47 | 0.54 | 0.46 | 0.75 | 0.54 | 0.97 | 0.95 | 0.35 |
| 2 | 22 | 0.78 | 0.93 | 0.01 | 0.02 | 1.00 | 0.99 | 0.93 | 0.78 |
| 3 | 22 | 0.33 | 0.96 | 0.50 | 1.00 | 0.01 | 1.00 | 1.00 | 0.35 |
| 4 | 22 | 0.41 | 0.69 | 0.05 | 0.19 | 0.94 | 0.88 | 0.82 | 0.41 |
| 5 | 22 | 0.66 | 0.02 | 0.02 | 1.00 | 1.00 | 0.52 | 0.92 | 0.66 |

## 3. Feedback for Improving Learning and Teaching Practice

### *3.1 Feedback for Improving Learning*

There is a lot of evidence showing that self-regulated learners are more persistent, resourceful, confident and tend to be higher achievers (Zimmerman & Shunk, 2001). In self-regulated learning, students undergo a recursive triptych: (1) forethought phase—analyze tasks, set goals, and plan behaviors before learning efforts; (2) performance phase—monitor and control their behaviors during learning efforts, and (3) self-reflection phase—self-evaluation based on feedback after learning efforts. Self-reflection processes in turn affect forethought processes and beliefs regarding subsequent learning efforts, which complete the self-regulatory cycle (Zimmerman, 2007). Intelligent self-regulation requires that students have in mind some goals to be achieved (Nicol & Macfarlane-Dick, 2006). However, students, particularly novices who lack prior knowledge of learning tasks, have

difficulty in setting goals and recognizing what the task should be done next (Kostons, van Gog, & Paas, 2012). For example, reactive learners, who self-regulate less effectively, set vague goals and as a result fail to evaluate (Reeve, Ryan, Deci, & Jang, 2008). Hence, assessment is needed to facilitate the development of self-regulation in students. In particular, CDA can contribute to foster self-regulation.

Nicol & Macfarlane-Dick (2006) proposes seven principles of good feedback practice in relation to the development of self-regulated learning. As one of them, Nicol & Macfarlane-Dick (2006) indicates the importance of clarifying assessment goals and standards. Indeed, some researches have already demonstrated that these approaches are effective (Suzuki, in press; Weaver, 2006). Suzuki (in press) showed that a rubric was helpful in informing students about the test's goals, grading standards, and their achievement within those standards. The students who received a rubric had higher intrinsic motivation and used more deep-processing strategies than those who did not receive it. In CDA, attributes correspond to learning goals and attribute mastery probabilities do to students' achievement within standards. Therefore, CDA can be used to clarify goals, analyze cognitive processing and knowledge state required in solving the problem, and provide effective feedback.

### *3.2 Feedback for Improving Teaching Practice*

CDA can provide not only usable information that helps students improve their learning, but also good information to teachers. That is, each individual student's attribute mastery probabilities are useful information for teachers to know about each student's detailed knowledge state and give them appropriate guidance individually for their future remedial work. Moreover, 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 have difficulty to know and deal with student's problem individually. So, we also performed a hierarchical cluster analysis using Ward's method to categorize students based on students' attribute mastery probabilities. As a result, 144 students were classified into four groups. Table 3 shows the four clusters' profiles, the sizes and mean scores of each cluster, which could be useful for teachers to adjust their teaching contents and activities in group guidance.

Table 3: 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 |

It is clear that students in the 1st cluster have mastered all attributes completely, which indicates about 67% of all the students in our study have achieved the level required by "Japanese government curriculum guidelines". While students in the 2nd 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 3rd 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.

Especially students in the 4th 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 which are the easier skills supposed to be acquired in 3rd grade. In fact, even students in the 3rd cluster, which have the lowest average score among the four clusters, have mastered the skill A5 almost perfectly. The students in the 4th cluster might have conducted addition and subtraction under a wrong rule: for example applying rules of multiplying or dividing for adding or subtracting (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.

### *3.3 Reporting Diagnostic Results by Descriptive Documents*

As shown above, CDA has the potential to guide students in their learning and teachers in their teaching. The success of providing effective feedback depends, in part, on method how to transmit the information. In fact, attribute mastery probability is a difficult concept and hard to grasp for both students and teachers because the meaning of probability is usually difficult to be understood. For effective feedback, assessment results need to be provided in the form of written documents containing statement that describe assessment standard and students' achievement. Needless to say, the descriptions should be written clearly and simply (Shute, 2008). However, how to effectively communicate such complex and detailed information on CDA specifically, has not been adequately studied yet (Robert & Gierl, 2010). In this study, an approach for reporting diagnostic results is proposed and implemented in a web-based system.

We provide rank categories for performance levels as an aid to interpreting attribute mastery probabilities. Of course, it is hard to say how many categories or how large each category are appropriate and should be, but for demonstration we use three rank categories as follows in this paper: "Excellent" for the range of mastery probability from 80% to 100%, "Good" for those from 50% to 79%, "Failed to achieve" for those from 0% to 49%. Even if students understand their level of achievement, they might not be motivated to learn further if they have no idea how to improve their performance. In our system we not only provide students with detailed information about what cognitive attributes were measured by the test and the degree to which the students have mastered these attributes, but also show them possible approaches to enhance their comprehension. Figure 1 shows an example of descriptive score reports on Moodle for a student regarding his two attributes' mastery out of eight. The report includes the degree of attainment, diagnostic information and guidance for the student to improve his learning, and skill description. Students can grasp their own achievement within standards by referring to the reports. If they have not mastered the skills yet, they can obtain information to fill the gap with learning goals from. Moreover, students and teachers can refer to each skill description, which informs them the learning goal to be achieve, by moving the mouse over each attribute's name.

In addition to the diagnostic feedback information about individual students, formative group profiles derived from cluster analysis can be useful for teachers as well to adjust their teaching contents and activities in the classroom. Figure 2 shows an example of descriptive score reports about one group profile, including characteristics of performance and guidance for improving teaching practice for each group. The radar chart of attribute mastery probabilities for each group gives a more visual display of overall mastery status for teachers to have a clear grasp and understanding about the class overall.

## 4. Web-based Assessment System

To provide fine-grained knowledge of students' performance, a web-based assessment clearly has potential advantages (e.g., Gikandi, Morrow, & Davis, 2011; Wang, 2011). One of them is its accessibility and timeliness. Students can take cognitive diagnostic tests on the web and get feedback immediately. Teachers can have immediate access to individual students' diagnostic information, group profiles, or information about test items they conducted. Additionally, web-based environments have capability to manage and organize large amount of information by using tools such as widget, which are not available in a static format on paper. For example, we can create tooltips to display diagnostic reports for both students and teachers for further information. Say when putting the mouse pointer over the attribute "changing fraction forms" in Figure 1, a tooltip will appear to show the

details about its description. We can also use widgets to provide further information for teachers from each cluster to

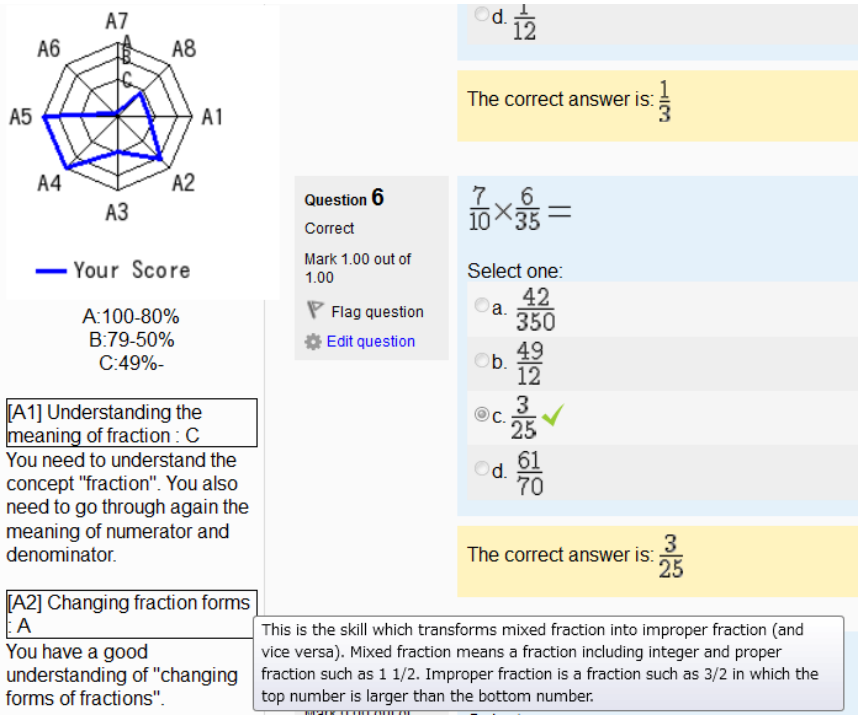


Figure 1. Sample diagnostic feedback for individual students

refer to individual diagnostic information of the students classified in the cluster.

As a platform, we used Moodle, a Learning Management System (LMS) or Virtual Learning Environment (VLE) widely used for its usability and scalability, to develop our web-based assessment system for students and teachers to access through their web browsers. We implemented items of a cognitive diagnostic fraction test in multiple-choice form on the Moodle platform. Students are asked to solve the problems on the web and after they finish the test, they will get all the diagnostic information feedback immediately just as stated above. The basic analyses are performed by using Moodle's modules, and for CDA analysis because there is no existed module we implemented DINA model in R, which is an open source statistics language, into the system by linking the R module through PHP program to Moodle platform.

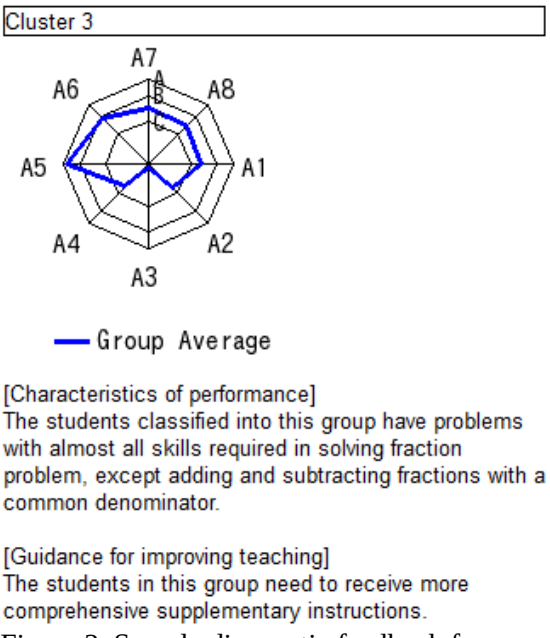


Figure 2. Sample diagnostic feedback for teachers with group profiles

## 5. Summary and Discussion

How to assist students and teachers to improve learning and teaching practice is a very important issue in education. In this study, we reviewed the concept of effective feedback for both students and teachers, and suggested applying CDA for feedback. To undergo successful self-regulated learning, students need to know learning goals, discrepancies between required and actual performance, and guidance for improving learning. We showed that CDA is applicable to provide effective feedback, and we presented an approach for reporting diagnostic information for CDA. Group-level diagnostic information in addition to individual student-level diagnostic information is also considered useful for teachers to grasp situation about the whole class and adjust their teaching contents and activities in

group guidance. We proposed classification schemes to categorize students into different groups based on their knowledge states. Furthermore, we implemented a web-based CDA system on Moodle platform, which can expand the potential of CDA through immediate analyses and linkages to large amount of diagnostic information for more effective feedback.

There are two directions for our future research. Firstly, we will examine the effects of our feedback approach in practical situation. This paper focused on design for effectively reporting CDA results, but obviously how students and teachers interpret and use feedback information is another key point. If they misinterpret the information we provide or cannot translate guideline for improvement into action, we need to revise our method for displaying diagnostic results. At mean time, construction and development of a web-based assessment system is another goal of our study. How to implement different functions to the system and make them easily usable for different users is a crucial issue and need to be further investigated and verified.

## References


American Federation of Teachers, National Council on Measurement in Education, National Education Association (AFT, NCME, NEA) (1990). Standards for teacher competence in the educational assessment of students. *Educational Measurement: Issues and Practice, 9*(4), 30-32.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review, 24*(2), 205-249.

DeLuca, C., & Klinger, D. (2010). Assessment literacy development: identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*(4), 419-438.

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education, 57*(4), 2333-2351.

Higgins, R., Hartley, P., & Skelton, A. (2001). Getting the message across: The problem of communicating assessment feedback. *Teaching in Higher Education, 6*(2), 269-274.

Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*(3), 258-272.

Kostons, D., van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction, 22*(2), 121-132.

Lee, Y., & Sawaki, Y. (2009). Application of three cognitive diagnosis models to ESL reading and listening assessments. *Language Assessment Quarterly, 6*(3), 239-263.

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research, 95*(4), 203-213.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Reeve, J., Ryan, R. M., Deci, E. L., & Jang, H. (2008). Understanding and promoting autonomous self-regulation: A self-determination theory perspective, In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and Self-Regulated learning* (pp.223-244). New Jersey: Lawrence Erlbaum Associates.

Roberts, M., & Gierl, M. (2010). Developing score reports for cognitive diagnostic assessments. *Educational Measurement: Issues and Practice, 29*(3), 25-38.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Suzuki, M. (in press). Exploring the mechanism of the effects of presenting a rubric and fact patterns. *Educational Technology Research, 36*.

Takahashi, T., Sun, Y., & Kakinuma, S. (2011). Development of attributes of cognitive diagnostic test in fraction calculations, *proceeding of 2011 conference of the Japan Association for Research on Testing* (pp.220-221).

Tatsuoka, K. (1985). A probabilistic model for diagnosing misconceptions in the pattern classification approach. *Journal of Educational Statistics, 10*(1), 55-73.

Wang, T. (2011). Developing web-based assessment strategies for facilitating junior high school students to perform self-regulated learning in an e-Learning environment. *Computers & Education, 57*(2), 1801-1812.

Weaver, M. (2006). Do students value feedback? Student perceptions of tutor's written responses. *Assessment & Evaluation in Higher Education, 31*(3), 379-394.

Zimmerman, B. J. (2007). Goal setting: A key proactive source of academic self-regulation. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp.267-295). New Jersey: Lawrence Erlbaum Associates.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives.* New Jersey: Lawrence Erlbaum Associates.