Exploring the Mechanism of the Effects of Presenting a Rubric and Fact Patterns*

Masayuki SUZUKI*1

*1 National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430 Japan

Received for publication, September 21, 2012

In the present study, we investigated the mechanism of the effects of presenting a rubric and fact patterns corresponding to each grading standard. 95 eighth-grade students were randomly assigned to 1 of 3 classrooms, each of which corresponded to an experimental condition. The participants studied mathematics for 5 sessions: at the end of the first day and the third day, the teacher gave a test on that day’s lecture. The feedback methods for the test results varied according to each classroom’s experimental condition: the rubric condition, the rubric plus fact pattern condition, or the comment-only condition. The results showed that students who received the rubric were more likely to consider that the purpose of the test was to show improvement than those who did not receive the rubric, and the rubric influenced intrinsic motivation, learning strategies, and test scores through values of a test. In addition, the results of this study suggested that informed assessment had an important role in values of test change. The results also revealed that fact patterns had no effect on the dependent variables.

Key words: informed assessment, rubric, values of a test, motivation, learning strategy

1. INTRODUCTION

Over the years, various methods have been devised to assess students’ academic achievement (e.g., Kojima and Iwaya 2010). However, tests organized by teachers, particularly evaluation through regular examinations, seem to play an important role in assessing student achievement. Especially in Japan, test performances are important for employment examinations as well as entrance examinations at senior high schools or colleges. Because such tests are society’s measures for screening applicants (Nagao 2000), test performance is highly important for students. It is essential, therefore, to identify the effects of a test and propose procedures for effectively conducting a test.

Academic achievement tests are expected to encourage planning, enhance learning motivation, and improve learning. Much research, however, has indicated a negative aspect of conducting tests: “the instruction of a test” (experimenters informed participants that their performance was being used to assess their competency) and “the instruction of grading” (experimenters informed participants that they would be graded based on the result of assessment) decreased intrinsic motivation (e.g., Maehr and Stallings 1972; Sakurai 1989). In addition, not many students regard assessment feedback as useful or as an opportunity to review and improve their learning. In fact, students’ tendency to learn for a short-term test and use superficial and rote learning has become a problem (e.g., Fujisawa 2002; Gipps 1994; Macelllan 2001; Nagao 2000). Thus, a test can adversely affect learning motivation and learning strategies. Yet, previous research has not examined a procedure for conducting a test that effectively enhances motivation and fosters effective learning strategies.

In recent years, researchers have focused on student perspectives concerning tests and assessments. They have shown that the effects of tests and assessments depend on students’ perceptions of those tests and assessments (e.g., Brown and Hirschfeld 2008; Peterson and Irving 2008; Suzuki 2009). For example, students’ perceptions such as “A test is irrelevant to learning” and “A test is conducted to force students to learn” weaken intrinsic motivation. In contrast, perceptions such as “A test helps to create a learning plan” and “A test helps to
improve learning” enhance learning motivation and foster effective learning strategies (e.g., Suzuki 2009). In addition, forming appropriate perceptions of a test in students can contribute to positive effects of the test. Furthermore, some studies have shown that an informed assessment approach1) (informing students about the purpose of a test and grading standard) changes students’ perception of the test and has a positive effect on motivation and strategies (e.g., Murayama 2006a; Suzuki 2011).

For example, Suzuki (2011) examined the effects of presenting a rubric to participants. A rubric is an “evaluation indicator which contains a scale of possible points and descriptions about writing and performance for each level of scale” (Nishioka 2003). It is expected to help students understand a test’s grading standards and judge their understanding level and to act as a guide to learning and self-evaluation (Tanaka 2008). For example, a rubric enables students who received Grade 2 to understand why they received this grade and how they can get a higher grade. Suzuki (2011) investigated the effects of instruction about a test’s grading standard and a guide on improving student achievement by presenting a rubric in an experimental mathematics class. The results showed that participants who received the rubric were more likely to consider that the purpose of the test was to facilitate improvement than those who did not receive the rubric but the comments on why their answers were incorrect and how they could be corrected. Furthermore, a path analysis indicated that the rubric influenced learning motivation, learning strategies, and test scores via values of a test. Suzuki (2011) suggested that the effects of a test on motivation and strategies depended on the value that students put on the test. Thus, the approach to informed assessment such as presenting a rubric had an important role in forming the appropriate values of a test. Suzuki’s research, however, has two problems.

First, Suzuki (2011) did not substantiate the mechanisms by which presenting a rubric helped students to regard the test as being important for improvement. In other words, although Suzuki found that presenting a rubric affects students’ perceptions of a test, he did not investigate whether the effects were caused by the informed assessment approach. Second, because the form of a rubric varies according to the purpose of the test (Mashiko 2003) and because no unique statement of the rubric is identified, we must examine what types of rubric introduce what effects. Therefore, it is important to study what types of rubrics are effective when introducing a rubric into educational practices.

Considering these problems, the present study investigates the effects of presenting a rubric in an experimental mathematics class, as did Suzuki (2011). To be more precise, we investigate whether a rubric has an effect on values of a test via the attainment of informed assessment. In addition, we focus on concreteness of grading standards and examine the effects of presenting fact patterns (example answers produced by students) corresponding to each grading standard along with the rubric. Presenting such fact patterns is effective in informing students about the qualitative differences between grades (e.g., Nishioka 2003). Because previous research has not investigated the effects of presenting fact patterns, examining this aspect of test assessment is of practical importance.

2. METHODS

2.1. Participants

Ninety-five 8th grade students participated voluntarily in the experimental classes called learning seminars. The five-day experimental classes began on August 2 and were conducted on alternate days. The participants were recruited from public junior high schools in Bunkyo and Taito wards and a junior high school affiliated with the University of Tokyo. They were randomly assigned to one of three classrooms, each of which corresponded to one of our three experimental condition: the comment-only condition, in which 32 participants received comments by the experimenters (17 males, 15 females); the rubric condition, in which 31 participants received the rubric (14 males, 17 females); and the rubric-plus-fact-pattern condition, in which 32 participants received both the rubric and fact patterns (17 males, 15 females).

2.2. Procedure

2.2.1. Pre-survey

One month before the classes started, the participants completed a pre-survey via a mailed questionnaire. The format for all items in the pre-survey was a 5-point scale, ranging from 1 = “not true” to 5 = “true.” Furthermore, to reduce the burden on the participants, the number of items was restricted to the minimum required.

Achievement goals Items developed by Tanaka and Yamauchi (2000) were used to assess
achievement goal orientations for mathematics, (1) mastery goal: focus on task mastery (three items), (2) performance–approach goal: focus on demonstrating competence relative to others (three items).

**Interest in mathematics** Items developed by Suzuki (2011) were used to assess students’ interest in mathematics (three items).

**Learning strategies during lessons** Items developed by Suzuki (2011) were used to assess learning strategies that students use during mathematics lessons in their schools. These included (1) deep-processing strategies that involve grasping the meaning of solutions (three items) and (2) surface-processing strategies that involve simply learning the solutions by heart (three items).

### 2.2.2. Lessons

Each lesson was held in the university by one of the authors. The classrooms for all three groups were identical. Four lessons were held on alternate days. In conducting experimental classes, the experimenter made sure that time allocation, lesson content, and content written on the blackboard were kept consistent across the conditions. Each lesson lasted 50 minutes. The participants listened to a lesson on linear equations with two unknowns (lesson topics were “number of pieces,” “speed,” “quotient,” and “concentration”). All participants had previously studied linear equations with two unknowns in their schools. Confirmation tests on lessons of that day and the previous day were conducted at the end of the first and third days, and the feedback was provided at the beginning of the next lessons. Reviewing the previous day’s test was allowed after the participants received the feedback. The experimenters collected the participants’ notebooks after each lesson (to control learning time) and returned them at the beginning of the next lesson. On the last day, the participants reviewed the lesson for 10 minutes and then completed a post-test.

#### 2.2.3. Confirmation test

Word problems on linear equations with two unknowns were administered. The first day’s test consisted of two problems, and the third day’s test included three problems. One question was an isomorphic problem with the same solution but different cover stories and numbers from the problems provided in the lessons; the other questions were challenging problems related to the lesson topic. The experimenters informed the participants that the tests’ purpose was to assess their understanding of the subject and told them to complete as many as possible. Confirmation tests lasted 10 minutes on the first day and 15 minutes on the third day.

#### 2.2.4. Feedback

The experimenters decided whether each answer was correct or not, graded and then returned each participant’s test in their notes. The method of providing feedback with test results varied according to each class’s experimental condition. In the comment–only condition, the participants received comments on why their answers were incorrect and information on how they could correct them. In the two rubric conditions, the participants received no comments but the rubric (Table 1) at the beginning of the second day’s lesson. The experimenters explained the purpose for presenting the rubric and the process of setting grading standard as follows: (1) grading standard was set with reference to the answer sheets, (2) grading was designed not to rank or compare the participants but to show their level of understanding, and (3) it is essential to

<table>
<thead>
<tr>
<th>Table 1. Rubric presented to the participants in the two rubric conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Shows two correct expressions and answers</td>
</tr>
<tr>
<td>4 Shows two correct expressions but no answer</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1 Shows expressions, figures, or tables which are imperfect but partially correct</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
judge their own understanding level and determine the best approaches for correcting mistakes in order to enhance their comprehension or gain greater achievement. In the rubric-plus-fact-pattern condition, the participants also received fact patterns that corresponded to the scores. The fact patterns were organized by the experimenters with reference to the answer sheets.

Additionally, all participants received their own answer sheets, sample answers, and points for reviewing the test. The sample answers contained both typical and different solutions.

2.2.5. **The instruction of reviewing**

All participants were instructed about learning strategies for reviewing a test on the second day. Initially, the experimenters distributed a sample way of reviewing and asked a few participants to state the characteristics of the sample. The experimenters then explained effective reviewing as follows: (1) review a problem and think of better approaches irrespective of correctness, (2) draw lessons from failures, and (3) summarize the points in their own words.

2.2.6. **Post-test**

The post-test consisted of two types of problems: (1) lesson isomorphic problems (three items), which had the same solution but different cover stories and numbers from the problems provided only in the lessons; (2) confirmation test isomorphic problems (three items), which had the same solution but different cover stories and numbers from the problems provided only in the confirmation tests (i.e., challenging problems). The participants were instructed to solve their first item in three minutes and the others in twenty-five minutes (i.e., five minutes per item). In addition, they were told to turn a page on cue and not to turn another page until after another cue even if they finished solving a problem. The experimenters also told the participants to complete as many problems as possible.

2.2.7. **Dependent variables**

**Degree of attainment of the informed assessment** This scale assessed the degree of attainment of the informed assessment. It comprised three sub-scales, and the format for all items was a 5-point scale. The measurement was conducted at the end of the second day's lesson. The sub-scales were the following: (1) Understanding, which measured the degree of students' understanding of the test's purpose being to judge their level of achievement and improve learning (three items); (2) Clarity, which measured the degree of the clarity of grading standard and improvement strategies (seven items); and (3) Utility, which measured the utility of reviewing a test. The participants rated eight strategies selected from strategies in reviewing a test scale (Suzuki 2010) as to perceived utility.

**Values of a test** Items developed by Suzuki (2011) were used to assess the "values of a test in the experimental class," which was defined as the participants' perception of the post-test. The format for the items was a 5-point scale. The measurement was conducted at the end of the fourth day and included the following: (1) Improvement, which helped to understand the extent of their own understanding and improve learning (six items); (2) Comparison, which helped teachers to compare the participants (five items); and (3) Enforcement, which forced students to learn (three items).

**Intrinsic motivation** Items developed by Suzuki (2011) were used to assess intrinsic motivation for lessons. The format for the items was a 5-point scale and the scale comprised four sub-scales, each having three items. The measurement was done after the post-test on the fifth day and included the following: (1) Enjoyment, assessing the degree of enjoyment or fun during lessons; (2) Challenge, assessing the degree of effort students exerted to get better achievement; (3) Autonomy, assessing the degree of perceiving that their approach to learning is not under control of teachers but autonomous; and (4) Curiosity, assessing the degree of interest in and delving into lesson topic.

**Learning strategies for a test** Learning strategies for reviewing lessons before the post-test on the fifth day were assessed after the post-test. We call these strategies "learning strategies for a test." The format for the items was a 5-point scale and the scale comprised deep-processing and surface-processing strategies (each scale has six items).

**Post-test** During the analysis, we computed student test scores on the basis of whether or not an item was answered correctly. If the expression was correct, the score was 1; if expression was incorrect, the score was 0.

3. **RESULT**

The analysis of each questionnaire was
conducted for the participants who attended on the day when the questionnaire was completed. Test scores were analyzed for the participants with perfect attendance. We excluded Enforcement from the analysis because it lacks internal consistency (α = .52). Furthermore, we combined Understanding with Clarity because they were highly correlated (r = .71). SPSS 16.0 and Amos 19.0 were used for the analysis.

3.1. Covariates
Achievement goals were used as covariates in analyzing the values of a test. Interest in mathematics was used for intrinsic motivation, learning strategies during lessons was used for learning strategies for a test, and confirmation test scores were used for the post-test scores. An analysis of variance showed that no significant differences existed in the covariates (Table 2).

3.2. Dependent variables
A contrast analysis was used to test the effects of presenting a rubric and those of presenting fact patterns in two rubric conditions. Two orthogonal contrasts were constructed. The first contrast tested for differences between the two rubric conditions and the comment-only condition (rubric contrast: rubric [1], rubric-plus-fact-pattern [1], comment-only [-2]). The second contrast tested for differences between the rubric condition and the rubric-plus-fact-pattern condition (fact pattern contrast: rubric [-1], rubric-plus-fact-pattern [1], comment-only [0]). Because SPSS 16.0 does not allow us to assign a contrast coefficient when we add a covariate, we used the Helmert method\(^2\). Means, standard deviations, and a summary of the test of each dependent variable are presented in Table 3.

3.2.1. Informed assessment
The rubric contrast for Understanding/Clarity was significant at \( p < .05 \), and that for Utility was significant at \( p < .10 \). That is, the participants who received the rubric were more likely to fully understand the purpose of the test, the grading standard and the best approaches for correcting mistakes and improving performance, and perceive reviewing a test useful than those who did not receive it. Fact pattern contrasts, however, were not significant.

3.2.2. Values of a test
The contrast analysis showed that the rubric contrast for Improvement using a mastery goal covariate (the Helmert method) was significant at \( p < .10 \). That is, the participants who received the rubric were more likely to regard the purpose of the test to be to facilitate improvement. The rubric contrast for Comparison using performance-approach goal covariate (the Helmert method) was not significant. Fact pattern contrasts were also not significant.

3.2.3. Intrinsic motivation
The rubric contrasts for both Enjoyment and Autonomy (the Helmert method) were significant at \( p < .05 \) and that for Curiosity was significant at \( p < .10 \). That is, the participants who received the rubric were more likely to enjoy mathematics lessons and perceive their autonomy for the

Table 2. Means (standard deviations) of covariates in each experimental condition and summary of analysis of variance

<table>
<thead>
<tr>
<th>Experimental conditions</th>
<th>Comment-only ((n = 29, 28))</th>
<th>Rubric ((n = 29, 27))</th>
<th>Rubric plus fact pattern ((n = 29, 24))</th>
<th>(F)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest in mathematics ((\alpha = .78))</td>
<td>2.97 (1.06)</td>
<td>3.25 (1.14)</td>
<td>3.54 (1.10)</td>
<td>1.91</td>
</tr>
<tr>
<td>Mastery goal ((\alpha = .88))</td>
<td>3.48 (0.96)</td>
<td>3.90 (0.91)</td>
<td>3.75 (0.80)</td>
<td>1.27</td>
</tr>
<tr>
<td>Performance approach goal ((\alpha = .70))</td>
<td>3.34 (0.95)</td>
<td>3.20 (1.09)</td>
<td>3.37 (1.04)</td>
<td>0.30</td>
</tr>
<tr>
<td>Deep-processing strategies ((\alpha = .82))</td>
<td>3.45 (0.81)</td>
<td>3.69 (1.09)</td>
<td>3.68 (1.04)</td>
<td>0.55</td>
</tr>
<tr>
<td>Surface-processing strategies ((\alpha = .70))</td>
<td>2.54 (0.82)</td>
<td>2.38 (0.95)</td>
<td>2.41 (0.84)</td>
<td>0.29</td>
</tr>
<tr>
<td>Confirmation test</td>
<td>2.64 (1.06)</td>
<td>2.52 (1.16)</td>
<td>2.92 (1.10)</td>
<td>0.85</td>
</tr>
</tbody>
</table>
lessons. Fact pattern contrasts, however, were not significant.

3.2.4. Learning strategies

The rubric contrast for both deep-processing strategies and surface-processing strategies (the Helmert method) was significant ($p < .05$ and $p < .10$, respectively). That is, the participants who received the rubric were more likely to grasp the meaning of the solutions. In contrast, those who did not receive the rubric were more likely to learn solutions by heart. Fact pattern contrasts, however, were not significant.

3.2.5. Post-test

The rubric contrast for lesson isomorphic was (the Helmert method) significant at $p < .05$, implying that the participants who received the rubric achieved higher scores than those who did not receive the rubric. The rubric contrast for the confirmation test isomorphic was (the Helmert method) not significant. Fact pattern contrasts were also not significant.

3.3. Examination of process

We conducted a path analysis to re-examine the process found in Suzuki (2011). That is, we assumed that the rubric influenced learning motivation and strategies and test scores via values of a test. We excluded Comparison because it has little relation with the other variables. Additionally, to reduce the complexity of the model, we combined four sub-scales to create the intrinsic motivation value.

The model provided not enough of a good fit but an acceptable fit to the data: AGFI = .836, CFI = .945, RMSEA = .106 and SRMR = .094 (Figure 1). Tomarken and Waller (2003) argued that model fit can be poor because the omission of important variables, which in reality are implicated in the causal structure, contributes to discrepancies between the observed and implied covariance matrices. Because Enforcement, which was highly correlated with motivation and strategies in Suzuki (2011), was omitted from the model in this study, model fit might be poor.

3.4. Mechanism of effects on values of a test

The results of Suzuki (2011) were supported in this study. We, therefore, conducted a path analysis to investigate whether the rubric had an effect on the values of a test with the introduction

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>$\alpha$</th>
<th>Comment-only</th>
<th>Rubric</th>
<th>Rubric plus fact pattern</th>
<th>Contrast analysis t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$ $SD$</td>
<td>$M$ $SD$</td>
<td>$M$ $SD$</td>
<td>Rubric Fact patterns</td>
</tr>
<tr>
<td>Informed assessment</td>
<td>($n = 28, 29, 27$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding</td>
<td>.85</td>
<td>3.95 .67</td>
<td>4.29 .52</td>
<td>4.14 .50</td>
<td>2.06* .97</td>
</tr>
<tr>
<td>Utility</td>
<td>.87</td>
<td>3.91 .74</td>
<td>4.24 .62</td>
<td>4.11 .63</td>
<td>1.75* .74</td>
</tr>
<tr>
<td>Values of a test</td>
<td>($n = 29, 29, 29$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>.77</td>
<td>3.97 .60</td>
<td>4.37 .55</td>
<td>4.11 .54</td>
<td>1.69* 1.62</td>
</tr>
<tr>
<td>Comparison</td>
<td>.72</td>
<td>2.98 .91</td>
<td>2.59 .84</td>
<td>2.83 .69</td>
<td>1.38 1.00</td>
</tr>
<tr>
<td>Intrinsic motivation</td>
<td>($n = 28, 29, 27$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enjoyment</td>
<td>.82</td>
<td>3.83 .99</td>
<td>4.12 .88</td>
<td>4.01 .94</td>
<td>2.68** 1.03</td>
</tr>
<tr>
<td>Challenge</td>
<td>.86</td>
<td>2.54 1.05</td>
<td>3.00 1.13</td>
<td>3.17 1.36</td>
<td>1.15 0.05</td>
</tr>
<tr>
<td>Autonomy</td>
<td>.71</td>
<td>3.44 .67</td>
<td>3.95 .63</td>
<td>3.79 .73</td>
<td>2.29* 1.29</td>
</tr>
<tr>
<td>Curiosity</td>
<td>.82</td>
<td>3.50 .90</td>
<td>4.02 .78</td>
<td>3.93 .94</td>
<td>1.76* 0.94</td>
</tr>
<tr>
<td>Learning strategies</td>
<td>($n = 28, 29, 27$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep-processing strategies</td>
<td>.84</td>
<td>3.57 .71</td>
<td>3.94 .61</td>
<td>3.84 .56</td>
<td>2.05* 0.55</td>
</tr>
<tr>
<td>Surface-processing strategies</td>
<td>.75</td>
<td>2.55 .62</td>
<td>2.18 .63</td>
<td>2.31 .78</td>
<td>1.90* 0.76</td>
</tr>
<tr>
<td>Post-test</td>
<td>($n = 28, 27, 24$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lesson isomorphic</td>
<td>1.29</td>
<td>0.76</td>
<td>1.52 0.70</td>
<td>1.71 0.89</td>
<td>2.27* 0.15</td>
</tr>
<tr>
<td>Confirmation isomorphic</td>
<td>1.60</td>
<td>0.99</td>
<td>1.63 1.31</td>
<td>1.67 1.01</td>
<td>0.04 0.92</td>
</tr>
</tbody>
</table>

* $p < .10$, $^* p < .05$, $** p < .01$
of informed assessment. First, we examined the partial mediation model in which we assumed that the dummy variable for the rubric contrast influenced Improvement directly and through Understanding/Clarity and Utility. The results showed that Understanding and Utility influenced Improvement significantly, while the dummy variable did not ($b^* = .10, p = .29$). Second, we examined the complete mediation model in which we did not assume that the dummy variable influenced Improvement directly. The results showed that Understanding and Utility influenced Improvement significantly ($b^* = .23, p < .05; b^* = .39, p < .01$). Additionally, the fit of the complete mediation model ($AIC = 19.10, BIC = 40.98$) was better than that of the partial mediation model ($AIC = 20.00, BIC = 44.31$). The results suggest that the effect of presenting the rubric was mediated by Understanding/Clarity and Utility. Thus, the effectiveness of a rubric is contingent on the attainment of informed assessment.

4. DISCUSSION

4.1. Effects of presenting a rubric on values of a test

This study supported Suzuki (2011). Thus, the participants who received the rubric were more likely to regard the test's purpose to be to facilitate improvement. Additionally, the results showed that the participants who understood the test's purpose of assessing their understanding of the subject and indicating improvement, those who regarded the grading standards and improvement strategies as clear, and those who perceived reviewing a test to be useful were more likely to regard the test's purpose as facilitating improvement. These results confirm that the informed assessment approach played an important role in changing students' perceptions of a test. These results also suggest that even if teachers employ alternative assessment methods (such as performance assessment and portfolio assessment) or devise new methods, expected effects cannot be obtained if students do not understand "why teachers employ those assessments" or "what standards teachers grade students." Therefore, whatever assessment method teachers employ, an informed assessment approach is essential for the method to have a positive effect on students.

4.2. Relations between values of a test and learning motivation and strategies

This study also showed that conducting a test had no direct effect on learning motivation or strategies but an indirect effect via values of a test. Regarding the positive effects of Improvement on intrinsic motivation, the cognitive evaluation theory (e.g., Deci and Ryan 1985) explains that information–control distinction is one important determinant of the motivational impact of environmental events. In other words, because a test was perceived as an informational event for the participants, who regarded the test's purpose to be to facilitate improvement, intrinsic motivation was enhanced. Additionally, Improvement has a positive effect on deep–processing strategies and a negative effect on surface–processing strategies. Thus, participants who intended to not only know their results but also to assess their own understanding were likely to not simply learn solutions by heart but to grasp the meaning of the solutions both before and after testing.

These results suggest that changing values of a test contributes to enhancing motivation for learning and fostering effective learning strategies.
It is, therefore, important to conduct a test while keeping in mind how students value a test.

4.3. Effects of fact patterns

Presenting fact patterns was found to have no effect in this study. Presenting only a rubric might have been effective because a rubric for mathematics problems may be objective. However, in performance assessment such as an oral presentation, an art work, or an observation activity, it is difficult to clarify grading standards where a fact pattern will be effective, even if a rubric includes detailed contents.

4.4. Effects on test performance

Whereas Suzuki (2011) found a difference in the scores of isomorphic problems that were provided only in confirmation tests, the present study found a difference in the scores of isomorphic problems that were taught only in lessons. This contradictory result might be because of the differences in the research procedures. In Suzuki (2011), the confirmation test was conducted on the fourth day and participants took a post-test after receiving feedback and reviewing the confirmation test. In this study, on the other hand, the confirmation test was not conducted on the fourth day and participants took a post-test after reviewing the previous lessons. Therefore, the difference in this study was possibly a result of strategies used in learning for the post-test on the fifth day, when the participants might have reviewed the problems provided in lesson.

4.5. Limitations and future directions

This study found the effect of presenting a rubric to students and its mechanism. We should investigate further what type of rubrics we should use in education practices. Other limitations of this study include a short intervention period. Furthermore, because participation in this study was voluntary, the participants (or their parents) might have higher motivation than average junior-high school students (or their parents). Thus, we must consider long-term intervention in a real-world classroom situation in order to resolve these problems.

NOTES

1) Informed assessment is defined as an assessment in which "teachers inform students about the purpose and standard of assessments and students agree to those" (Murayama 2006b).
2) The Helmert contrast in SPSS 16.0 compares each level of the factor with the mean of succeeding levels. In this case, the two contrasts are [1, -1/2, -1/2] and [0, 1, -1].

ACKNOWLEDGEMENTS

We gratefully acknowledge the many helpful comments and suggestions we received from Professor Shinichi Ichikawa (the University of Tokyo). We are also grateful to our colleagues who supported this research.

REFERENCES


Exploring the Mechanism of the Effects of Presenting a Rubric and Fact Patterns

*Educational Psychology, 37*: 29-35 (in Japanese)


