An Improved Teaching Behavior Estimation Model from Student Evaluations

Yusuke Kometani[1], Takahito Tomoto[2], Takehiro Furuta[3], and Takako Akakura[2]

[1] Graduate School of Engineering
and
[2] Faculty of Engineering, Tokyo University of Science,
1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
{kometani,tomoto,akakura}@ms.kagu.tus.ac.jp
[3] Nara University of Education, Takabatake-cho, Nara City, Nara 630-8528, Japan
takef@nara-edu.ac.jp

Abstract. Many universities conduct student evaluations. Their purpose is to encourage improvement in teaching. However, the evaluations are merely subjective assessments by students, meaning that instructors cannot necessarily easily relate evaluations to areas for improvement in teaching. To address this issue, we suggest a teaching behavior estimation model that can estimate teaching behaviors from student evaluations of each lesson. In previous research, we built a model on the assumption that teaching behaviors are not correlated with other behaviors and that student evaluation items are uncorrelated to other evaluation items. However, this assumption could not be verified. Our research suggests a new teaching behavior estimation model that represents the correlation between factors of teaching and factors of student evaluations. To analyze this, we conducted canonical correlation between two kinds of factors and obtained correlations. This result shows that it is possible to construct a teaching behavior estimation model based on factors of teaching behavior and factors of student evaluations.

Keywords: Student evaluation, Lesson improvement, Teaching behavior, Teaching behavior estimation model.

1 Introduction

Student evaluations of teaching are a typical method for supporting lesson improvement by instructors. At present, student evaluations are used by many universities. However, these evaluations are subjective assessments by students, such as "a lesson is incomprehensible." Therefore, it is difficult for university instructors to discern concrete areas for improvement from student evaluations.

We previously proposed an approach that aims to support improvements in teaching [1]. We proposed a model (teaching behavior estimation model) that presumes teaching behavior can be measured from student evaluations and developed a function (teaching behavior estimation function) to identify teaching behaviors in the model

S. Yamamoto (Ed.): HIMI 2014, Part II, LNCS 8522, pp. 59–68, 2014.

that can be improved. In this research, analysis intended to improve the teaching behavior estimation model is conducted. Specifically, we increase the variety of teaching behaviors subjected to analysis and propose a model that yields the relation between the factor of teaching behaviors and the factors of student evaluations. The purpose of this research is to show that it is possible to build the model.

2 Related Works

Students' evaluation feedback methods and student reactions to support teaching improvement have been studied in the past. Stalmeijer et al. explored whether feedback effectiveness improved when physician teachers' self-assessments were added to written feedback based on student ratings [2]. The physician teachers considered the combination of self-assessment and student ratings more effective than either self-assessment or written feedback alone. The authors concluded that self-assessment can be useful in stimulating teaching improvement. However, there was no evidence that the teachers grasped points for improvement of teaching behaviors by reviewing students' evaluations. Thus, our proposed method additionally involves an objective evaluation of the teaching behavior based on evaluations obtained from the students and the data are then fed back to the teacher.

Recent years, estimation methods for students' evaluations based on teaching behaviors were studied. Large repositories of presentation recordings often provide users with rating facilities. Pietro et al. explored nonverbal behavior (in particular arms movement and prosody) allows one to predict whether a video presentation is rated as low or high in terms of quality[3]. The experiments have focused on the nonverbal behaviors most important in an oral presentation, namely pose, gestures, movements and prosody. The results show that the mean pitch and position of arms allow one to predict whether a presentation is rated as high or low quality. However, this study didn't explore whether it is able to predict teaching behaviors based on students' evaluations.

Lessons can be improved by an improvement in course content and by an improvement in teaching behaviors. We previously proposed an instructor support method that aims to help instructors improve their own teaching behaviors [1]. We analyzed the relation between the average value of student evaluations of specific lessons and the number of times particular teaching behaviors occurred in the lesson. As a result, the relation between each lesson evaluation and teaching behavior was obtained, and it turned out that it is possible to estimate the number of times that specific teaching behaviors occurred by using lesson evaluations. Then, a model was built that shows the relation between the average value of each lesson evaluation and the number of times each teaching behavior occurred in the lesson. Additionally, a teaching behavior estimation function was proposed to estimate the number of times each teaching behavior had occurred. Figure 1 shows the teaching behavior estimation model built to evaluate the teaching behavior estimation function. The result of the evaluation experiment showed that university teachers would like to use the function and that the estimated teaching behaviors were effective in supporting lesson improvement. These results show that increasing the variety of teaching actions and refining the model leads to more effective lesson improvement support.

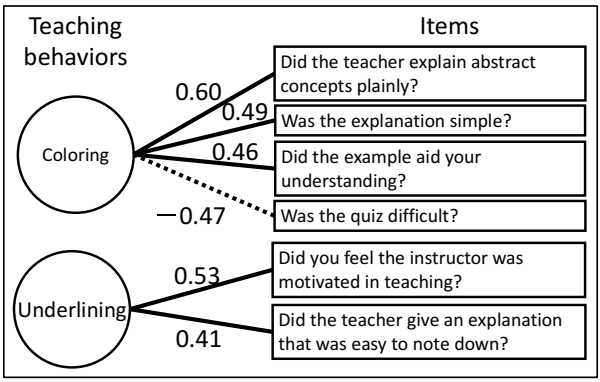

Fig. 1. Teaching behavior estimation model

3 Improvement of the Teaching Behavior Estimation Model

Among instructors who want to improve their lessons, the important teaching actions may differ by instructor. Moreover, according to the lesson form and course content, evaluation items may differ. In consideration of this, previous work built a model that assumed a one-to-one relation between lesson evaluation items and teaching behaviors (Fig. 1). Thus, the model in previous work assumes that teaching behaviors and student evaluation items are mostly uncorrelated. However, verification of this was not carried out in that work. If such correlations do in fact occur, then it is necessary to reinvestigate the correlation between the newly added teaching behaviors and the evaluation items and how the model is affected by the addition of new teaching behaviors. However, if a model is built that allows correlation between teaching behaviors or correlation between student evaluation items, it becomes possible to treat highly related mutual teaching behaviors equally. This allows minimal changes to the model with the addition of new teaching behaviors. Moreover, when estimating a certain teaching behavior, multi-correlation of student items can be prevented, which allows shorter evaluations. Therefore, when acquiring student evaluations, an advantage of this model type is that the burden on students is mitigated. For the above reasons, it is believed that an improved model should allow detailed analysis of the correlations among student evaluation items and among teaching behaviors.

This research aims at building an estimation model that is flexible on included teaching behaviors and student evaluation items. Figure 2 shows the model assumed by this research. The model comprises independent factors of teaching behavior and independent factors of student evaluation. Anything contained in a factor of an existing model can be incorporated into a factor into the revised model, but items not part of any existing factor require modification of the model to incorporation of the new factor.

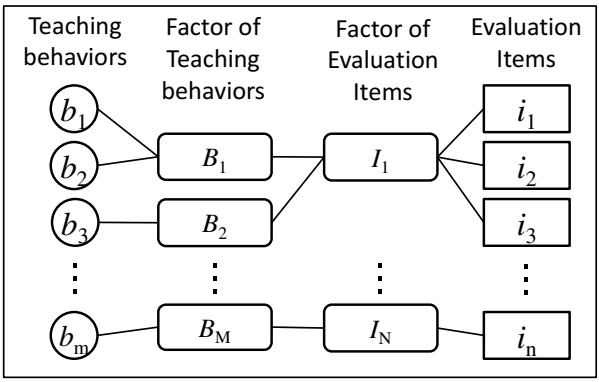

Fig. 2. The model incorporating factors

4 Methodology

In this research, the correlations between the independent factors of teaching behaviors and the independent factors of student evaluation are analyzed. First, data are collected in preparation. After determining the teaching behaviors to measure, collection of student evaluations and measurement of teaching actions are performed. Teaching actions in a lesson are measured by viewing and listening to video recordings of the lesson.

Then, the acquired data are analyzed. Factor analysis is applied to every teaching behavior and every student evaluation, which extracts the factors analytically. Canonical analysis is then performed on the factors of teaching actions and on the factors of student assessments, and correlations between factors are investigated.

4.1 Teaching Behaviors

We previously targeted two kinds of teaching actions, underlining and coloring [1]. However, to analyze the factors of teaching behaviors, it is necessary to measure many more kinds of teaching behaviors. We defined 27 concrete teaching behaviors to address this problem, on the basis of previous investigations into abstract teaching skill.

The behaviors defined in this research are blackboard writing behaviors, such as "uses an arrow or a line to connect two or more items written to the blackboard" and "draws a line on the blackboard and distinguishes the content when changing to a new topic"; oral behaviors, such as "orally checks whether there are any questions or comments from students" and "orally notes when changing to a new topic during the middle of a lesson"; and non-verbal behaviors, such as "folds arms when speaking" and "puts hand on waist when speaking."

4.2 Student Evaluations

The lesson is subject I at university T (a lesson about information mathematics). Student evaluations were acquired for 29 lessons (lessons 2–9, 11, 13 of the course in 2010, lessons 4–8, 10–13 in 2011, and lessons 3–9, 11–13 in 2012). The number of participants enrolled in the course varied from 80 to 110 students. To measure teaching behaviors, the lectures were videotaped. Table 1 shows some of the evaluation items. A 20-item questionnaire was used. A copy of the questionnaire was distributed at the end of each lesson. Participants answered each question item on a five-point Likert scale (from "1. Strongly disagree" to "5. Strongly agree"). The questionnaire required about 5 minutes to complete teach time. The questionnaire was completed in a machine-readable format, and students were informed in advance that the evaluation would not influence their course results. In addition, no feedback about the evaluation results was given to the instructor.

5 Results

The relations between the factors of teaching behaviors and the factors of student evaluations are analyzed. First, factor analysis is applied to student evaluations. The independent factors are extracted, and the features of each factor are considered. Next, factor analysis is applied to teaching behaviors, the independent factors are extracted, and the features of each factor are considered. Finally, canonical analysis is applied to the factors of student evaluations and the factors of teaching behaviors, and the relations among factors are clarified.

5.1 Measurement of the Frequency of Teaching Behaviors

The number of times that the instructor performed specified behaviors during each session was measured. Two fourth-year students viewed and listened to videos of lectures (about 2200 minutes) and measured the number of times that specific teaching behaviors occurred. To check the reliability of measurements, the two students' results were compared. Ten lecture videos and ten behaviors were randomly selected. When the two students were compared on number of recorded occurrences, the measurement results for 9 of 10 behaviors were in agreement. We treat this as sufficient reliability in measurement and decided to use the data for analysis. The measured teaching behaviors were totaled by lesson.

In 29 lessons of subject I, 23 kinds of teaching behaviors were observed. Among teaching behaviors, 6 were seen only once in 29 lessons. We therefore decided to use data on the remaining 17 kinds of teaching behaviors, excluding the teaching behaviors seen only once.

Table 1 shows the value of the number of times that the teaching behaviors were seen per session. The table shows that the number of times depends strongly on the kind of teaching behavior. Certain behaviors stand out as occurring particularly

frequently, such as “underlines a word or a sentence…,” “draws a frame…,” “changes the color of chalk…,” and “enumerates two or more items on blackboard.” The instructor tends to frequently use blackboard actions for emphasis.

Table 1. The value of the number of times teaching behaviors occur per lesson

Teaching behaviors	2010 2nd	2010 3rd	···	2012 13th
Underlines a word or a sentence on the blackboard	6	16	···	7
Draws a frame around two or more pieces of information on the blackboard	6	15	···	10
Changes the color of chalk when writing a word or a sentence on the blackboard	7	4	···	19
Enumerates two or more items on the blackboard	6	7	···	5
Uses an arrow or a line to connect two or more items written on the blackboard	10	5	···	7
Draws a frame around one word or sentence on the blackboard	0	0	···	13
Asks whether students understood the content	3	7	···	3
Set the task for students during the session	2	3	···	0

5.2 Factor Analysis

Factor analysis was applied to each evaluation item. Factors were extracted for eigenvalues greater than 1. As a result of factor analysis, three factors were extracted whose cumulative contribution ratio was 80.67%. Factor loadings are shown in Table 2. It is thought that the f_{i1} is “an evaluation factor about the lesson content,” because it has a large absolute value for the factor loadings “Did you understand the lesson content” “Did you master new knowledge?” and “Were you satisfied with the lesson?” It is thought that f_{i2} is “an evaluation factor about the method of methods of delivery” because it has a large absolute value for factor loadings “Was the writing on the blackboard legible?” and “Did the instructor give an explanation that was easy to note down?” It is thought that f_{i3} is “a factor of self-teaching” because it has a large absolute value for factor loadings “Did you prepare for the lesson?” and “Did you review the pre-lesson notes?”

After this, factor analysis was applied to teaching behaviors. Factors were extracted for eigenvalues more than 1. As a result of factor analysis, four factors were extracted whose cumulative contribution ratio was 39.42%. Table 3 shows the factor loadings. Factor f_{b1} is a “question-emphasis behavior factor” because it has a large absolute value for factor loadings “orally checks whether there are any questions or comments from students” and “orally notes when changing to a new topic during the middle of a lesson.” Factor f_{b2} is a “blackboard writing behavior factor” because it has a large absolute value for factor loadings “changes the color of chalk when writing a word or a sentence on the blackboard” and “uses an arrow or a line to connect two or more items written on the blackboard.” The factor f_{b3} is “planned explanation factor” because it has a large absolute value for factor loadings “enumerates two or more

items on the blackboard" and "explains a certain concept orally according to the phenomenon which may happen actually." The factor f_{b4} is an "overbearing behavior factor," because it has a large absolute value for factor loadings "folds arms when speaking" and "puts hand on waist when speaking."

Table 2. Factor loadings (student evaluations)

Evaluation Items	f_{i1}	f_{i2}	f_{i3}
Did you understand the lesson content?	**0.94**	0.08	0.12
Did you master new knowledge?	**0.91**	0.24	-0.02
Did you grasp the importance concepts of the lesson?	**0.90**	0.14	0.30
Were you satisfied with the lesson?	**0.88**	0.38	0.16
Were the explanations simple?	**0.88**	0.36	0.16
Were you interested in the lesson?	**0.84**	0.20	0.26
Did the examples aid your understanding?	**0.80**	0.51	0.03
Did the instructor plainly explain abstract concepts?	**0.78**	0.46	0.16
Was your attitude to study positive?	**0.64**	0.14	-0.05
Was the volume of the instructor's voice suitable?	0.07	**0.89**	-0.01
Was the writing on the blackboard legible?	0.26	**0.86**	-0.02
Did you feel the instructor was motivated in teaching?	0.25	**0.83**	-0.03
Did the instructor give an explanation that was easy to note down?	0.50	**0.80**	0.11
Was the quantity of material in the lesson appropriate?	0.49	**0.60**	0.36
Did you prepare for the lesson?	0.13	-0.14	**0.98**
Did you review the pre-lesson notes?	0.12	0.12	**0.89**

Table 3. Factor loadings (teaching behaviors)

Teaching behaviors	f_{b1}	f_{b2}	f_{b3}	f_{b4}
Orally checks whether there are any questions or comments from students	**0.76**	0.15	0.52	0.07
Orally notes when changing to a new topic during the middle of a lesson	**0.60**	0.26	0.14	0.03
Uses an arrow or a line to connect two or more items written on the blackboard	0.06	**0.75**	0.41	-0.15
Changes the color of the chalk when writing a word or a sentence on the blackboard	-0.09	**0.54**	0.20	0.05
Draws a frame around two or more pieces of information on the blackboard	-0.07	**0.45**	-0.37	0.13
Enumerates two or more items on the blackboard	0.09	0.15	**0.52**	-0.06
Orally explains a concept by giving actual examples	-0.55	0.03	**0.48**	0.36
Set the task for students during the session	0.14	-0.25	**0.47**	0.15
Draws a line on the blackboard and distinguishes the content when changing to a new topic	-0.21	-0.02	**0.46**	-0.24
Folds arms when speaking	-0.15	0.39	-0.33	**0.67**
Puts hand on waist when speaking	-0.18	-0.27	-0.01	**0.41**

5.3 Canonical Correlation Analysis

Canonical analysis was applied to the teaching behavior factors and the student evaluation factors by using each factor score for 29 lectures of subject I. Three axes were extracted by canonical analysis. Table 4 shows the structural coefficients. The considerations of each axis of canonical analysis are shown below.

Axis 1

- The absolute values of the structural coefficients of the "Lesson content" and "Deliberate explanation" are large. If there is a lot of planned explanation, the evaluation of the lesson content is likely to be low. This indicates that there is excessively planned behavior and too much content in the lesson (rushing through many topics by using itemized statement items, etc.), which interrupts student understanding of the content.

Axis 2

- The absolute value of the structural coefficient of "the evaluation factor about the method of delivery" and "the behavior factor of blackboard writing" is large. If blackboard writing behavior occurs often, the evaluation of the method of delivery will be high.

Axis 3

- The absolute value of the structural coefficient of an "overbearing behavior factor" and "the factor of self-study" is large. Overbearing behaviors reduce the evaluation of "self-study."

Table 4. Structured coefficients

factors	Axis 1	Axis 2	Axis 3
free variables (teaching behavior)			
f_{b1}: Question and emphasis	0.33	-0.40	-0.47
f_{b2}: Blackboard writing	-0.47	**-0.86**	0.19
f_{b3}: Deliberate explanation	**0.80**	-0.26	0.35
f_{b4}: Overbearing behavior	0.01	-0.07	**-0.80**
bound variables (students' evaluation)			
f_{i1}: Lesson content	**-0.89**	0.36	-0.28
f_{i2}: Method of delivery	-0.46	**-0.82**	0.35
f_{i3}: Self-study	-0.07	0.39	**0.92**

Table 5. Correlation between factors

	fi1: Lesson content	fi2: Method of delivery	fi3: Self-study
f_{b1}: Question and emphasis	-0.21	-0.01	-0.19
f_{b2}: Blackboard writing	0.13	**0.42**	-0.06
f_{b3}: Deliberate explanation	**-0.51**	-0.12	0.01
f_{b4}: Overbearing behavior	0.05	-0.05	-0.20

The correlation coefficients between each factor pair are shown in Table 5. The following factor combinations were founded to be correlated:

— (a) "Deliberate explanation" and "Lesson content,"
— (b) "Blackboard writing" and "Method of delivery," and
— (c) "Overbearing behavior" and "Self-study."

For relations (a) and (b), a correlation of medium degree was found, and a weak correlation was found for relation (c). Thus, correlations can be observed between factors of teaching behaviors and factors of student evaluations. This result suggests that a model can be constructed that estimates the factors of teaching behavior from the factors of student evaluations.

The factors of teaching action and the factors of student evaluations can be clarified, and the following advantages accrue from building a model using the factors.

- A factor can be replaced without influencing other factors in the model.
  The correlation coefficient of a model will change if teaching behaviors or student evaluation items are exchanged in the model because of the one-to-one correlation between teaching behavior and student evaluation items. However in our proposed model between factors, instructors can exchange teaching behaviors and student evaluation items without affecting the correlation between factors. For example, if an instructor wants to evaluate the effect of chalk colors, the value of "the factor of blackboard writing" can be calculated from the number of times that color was used, and the results can be compared by means of the value of the average value of "the evaluation factor about the method of delivery." Thus, it is possible to include teaching behaviors in a model without changing the correlation model, such as here without having a student evaluation item on chalk color.
- It is possible to reduce the number of student evaluation items.
  The burden on students from evaluations can be reduced. For example, it is possible to reduce the number of items to 6 by using each 2 items for each of the 3 extracted factors.

6 Conclusions and Future Work

This research aimed at the improvement of the teaching behavior estimation model to support improvement in instructors' lessons. A model comprising teaching behavior factors and student evaluation factors was proposed. To verify whether such a model is feasible, canonical analysis was applied to the factors of teaching behaviors and the factors of student evaluations. As a result, the factors of teaching behaviors and the factors of student evaluations were correlated. This result showed that model construction is possible.

As a future subject, we plan to construct a more flexible model. For lessons in different academic years and on different subjects, measurement of teaching behaviors and acquisition of student evaluations will be performed, and the accuracy of the model will be estimated. Furthermore, it would be useful to show the relation between

the factors of teaching behaviors and the factors of student evaluations to instructors. By knowing the teaching behavior factors, student evaluation factors, and their relations, it should be possible to identify teaching behavior groups that would affect student evaluations. We would like to investigate the effect on lesson improvement from sharing such information with instructors.

Acknowledgements. This research was partially supported by a Grant-in-Aid for Scientific Research (No. 24300291, 2012–2015) from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

1. Kometani, Y., Tomoto, T., Furuta, T., Akakura, T.: Video feedback system for teaching improvement using students' sequential and overall teaching evaluations. In: Yamamoto, S. (ed.) HCI 2013, Part III. LNCS, vol. 8018, pp. 79–88. Springer, Heidelberg (2013)
2. Stalmeijer, R.E., Dolmans, D.H.J.M., Wolfhagen, I.H.A.P., Peters, W.G., van Coppenolle, L., Scherpbier, A.J.J.A.: Combined student ratings and self-assessment provide useful feedback for clinical teachers. Adv. in Health Sci. Educ. 15, 315–328 (2010)
3. Salvagnini, P., Salamin, H., Cristani, M., Vinciarelli, A., Murino, V.: Learning How to Teach from "Videolectures": Automatic Prediction of Lecture Ratings Based on Teacher's Nonverbal Behavior. In: 3rd IEEE International Conference on Cognitive Infocommunications, pp. 2–5 (2012)