

Quantitative Evaluation Method Examination for Mock Classes in a Faculty Development Program

Satori Hachisuka¹, Kayoko Kurita², and Shin'ichi Warisawa³

¹Interfaculty Initiative in Information Studies, The University of Tokyo, Japan

²Graduate School of Education, The University of Tokyo, Japan

³Graduate School of Frontier Sciences, The University of Tokyo, Japan

ABSTRACT

This research focuses on the mock classes in the faculty development program to nurture university teachers. This research aims to establish a quantitative evaluation method for mock classes and examine the possibility. In this paper, we analyzed the data from mock classes conducted by 23 participants of the faculty development program who mainly aim to become teachers. First, based on the subjective evaluation results by the Likert scale currently used in the mock classes, multiple regression analysis was performed using the comprehensive evaluation of the mock classes as the objective variable and the 11 evaluation items as the explanatory variables. As a result, four items were extracted as variables included in the regression equation: clarity of objectives and goals, class flow, time allocation, and clarity of lecturer's explanation. The coefficient of determination of this multiple regression equation was 0.93, and the multiple correlation coefficient was 0.97. Next, in this paper, we attempted to quantify gestures which we considered the component of the "clarity of explanation" related to the delivery method among the multiple regression equation variables. We adopted OpenPose for motion analysis using image processing for the recorded video of the mock classes. As a result, of the 23 mock class participants, when comparing the higher and lower groups than the average value of overall subjective evaluation, the higher evaluated group showed a tendency of large value in the total amount of movement of both hands, the range and speed of both hands' motion, and face direction.

Keywords: Faculty development program, Mock classes, Motion analysis, Quantitative evaluation method

INTRODUCTION

Faculty development (FD) originated in Europe and American colleges to improve university teaching methods and ensure the quality of university teachers (Park, 1979). It plays an essential role in fulfilling the objectives of advancing higher education and strengthening international competitiveness (Froyd, 2005). FD has recently become as important at Japanese universities as in other countries, and establishing FDs at graduate schools and universities became compulsory in 2007 and 2008, respectively (Ministry of Education, Culture, Sports, Science and Technology (MEXT)). The core educational items of the FD program include course design, class design, teaching

material preparation, and mock classes to learn these elements practically (Kilic, 2010; Remesh, 2013; Ozeki, 2023). Some established evaluation methods exist for items other than mock classes, such as lesson design models and teaching techniques. On the other hand, we have yet to base quantitative evaluation methods on mock classes requiring real-time classroom adjustments. In particular, regarding class content delivery, the mainstream is to evaluate impressions based on body language and speaking style, and quantitative evaluation criteria need to be clarified (Murray, 2020). Therefore, there is a problem in that it is difficult to propose or provide guidance on specific improvement measures for body language. Consequently, we are currently investigating the possibility of establishing a quantitative evaluation method for mock classes.

In this paper, as a primary study of quantitative evaluation methods for simulated classes, we analyzed the following two points based on data related to actual mock classes. (1) What items are emphasized in the subjective evaluation of mock classes? (2) What is the relationship between the evaluation of mock lessons and gestures?

DATA COLLECTION OF MOCK CLASSES

The University of Tokyo has offered the University of Tokyo Future Faculty Program (FFP) since 2013 as a teaching ability improvement program for graduate students, postdoctoral fellows, and young faculty members who aim to become university teachers (UTokyo Faculty Development). This program consists of DAY 1 to DAY 8. In the first half (DAY 1 to DAY 5), participants will acquire knowledge about class design and course design, and in the latter two sessions (DAY 6 and DAY 7), participants conduct mock classes. In addition, in the final lesson (DAY 8), participants will be able to create a structured academic portfolio chart that embodies what they want to be as a university faculty member or what they would like to achieve when they become university teachers, both on an individual level and as a university. The goal is to foster improvements in the educational abilities of participants.

In this study, we used the results of the peer evaluations and recorded videos of the mock classes conducted on DAY 6 and DAY 7 of the program. Informed consent was obtained from the participants in advance for utilizing the peer evaluation results and video recording for this research with the approval of the Dean of the Graduate School of Education, the University of Tokyo. In this paper, we analyzed data from the second mock class (DAY 7). The reason for this is to control the level of the mock classes among participants. The FD program targeted this time includes program participants who are not yet university teachers and who are university teachers but have little or no experience in delivering lectures, so we targeted the second mock class (DAY 7). By doing so, we decided to receive feedback from the first mock class (DAY 6) and ensure that the participants had acquired knowledge of the essential elements necessary for the mock classes. Specifically, in this paper, mock class data from the University of Tokyo Future Faculty Program database was randomly extracted and analyzed. Note that this paper

dealt with data for one class of a program offered in 2017, which includes 23 participants.

Evaluation Sheet

In general, evaluation sheets for mock classes vary depending on the purpose and policy of each university's faculty development program. So, there are many forms of evaluation. Still, most evaluation sheets used a five-grade evaluation, including class management, interaction, and communication content (Murray, 2020). The evaluation sheet for the mock class that was the subject of this analysis consisted of 11 elemental evaluations, comprehensive evaluations, and free descriptions. The elemental evaluation of the lesson was on a five-point scale (1: Strongly Disagree to 5: Strongly Agree), and the comprehensive evaluation was on a 10-point Likert scale (1: not good at all to 10: very good). Eight to 13 participants individually evaluated one mock class by completing an evaluation sheet. The 11 items for elemental evaluation are as follows.

1. The objectives and goals of the class were clearly defined.
2. The purpose/goal and content were consistent.
3. The level setting was appropriate.
4. The class had a good flow.
5. I wanted to know more about the content of the class.
6. The time allocation was appropriate for the content.
7. I felt like I was participating in the class.
8. The teaching materials were well-developed.
9. The lecturer's explanation was easy to understand.
10. I felt the enthusiasm of the lecturer.
11. The lecturer's delivery was good.

Recorded Movie

We used a digital video camera to record the mock class from the back of the classroom. The composition of the video was that the participants conducting the mock class stood in front or on the side of the screen for projecting the lecture slides. Furthermore, the mock class lecturer was able to move freely in front of the screen and was also able to move their body and face without any restrictions. Figure 1 shows an overview of the captured video. The mock class lasted about 6 minutes per person, and the content was related to each participant's field of expertise, so the content differed for each person.

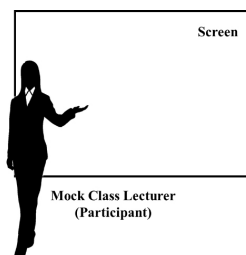


Figure 1: An example image of a mock class lecturer. The mock class lecturer stands in front of a screen and moves freely.

DATA ANALYSIS METHOD

The comprehensive evaluation of the mock class is related to the various elemental evaluations. Among these, we focused on the extent to which impressions about how the lesson content delivery was associated with the subjective overall assessment and how the size of body language movements was related to the comprehensive evaluation. Therefore, we investigated using multiple regression analysis and motion analysis. We can determine how mock classes should be improved by using multiple regression analysis to extract essential elements for evaluating mock classes and quantifying the communication methods that are considered difficult to quantify. The aim is to show which features are good and which components must be strengthened.

Evaluation Sheet

We collected the evaluation sheets filled out by audience participants for each mock class and tallied the Likert scale values. We gathered eight to 13 evaluation sheets for each lecturer and calculated the average score for each item. A multiple regression analysis was conducted on each participant's evaluation values using the mock lesson's overall evaluation as the objective variable and the 11 elemental evaluation items as explanatory variables. This analysis aims to understand how the elements of delivery style influence the comprehensive evaluation of the mock lesson. In addition, taking into account the fluctuation of each participant's interpretation in the subjective evaluation, among the items on the evaluation sheet mentioned above, three things as "9. The lecturer's explanation was easy to understand", "10. I felt the enthusiasm of the lecturer", and "11. The lecturer's delivery was good" were positioned as items related to delivery and way of explain methods.

Motion Analysis of Mock Class Lecturers

We utilized OpenPose for image analysis to measure the lecturer's movements during the mock classes. OpenPose is a motion analysis method using deep learning, which analyzes and outputs the two-dimensional coordinates of landmarks located on the virtual skeleton overlaid on the person for each video frame (Cao, 2021). The feature points on the virtual skeleton are placed at 25 points on the head, body, and limbs (Figure 2). Some studies of the uses of OpenPose in education include research on estimating students' concentration and interest levels from their movements and research on systems that train elementary and secondary education teachers to teach through interaction with CG students (Fukuda, 2019). However, there is no precedent for research using OpenPose in the field of FD.

In this paper, we focused on body language, considered an essential element of lecturer delivery, and quantitatively analyzed the movements of the lecturer's left and right hands and neck. The specific movement parameters were:

- i. The total amount of movement of both hands.
- ii. The range of motion of both hands.
- iii. The speed of movement of both hands.
- iv. The total amount of the face direction during the mock class.

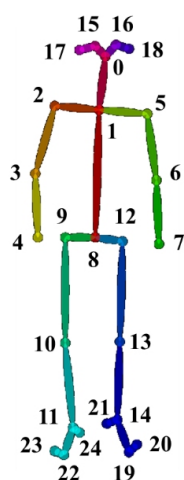


Figure 2: The virtual skeleton of openpose and 25 landmarks.

In addition, we normalized each index by dividing each parameter on the length between each lecturer's nose and neck landmarks to reduce the influence of differences in the size of human images recorded on video and individual differences in body size. For each parameter, we calculated the value for each second based on the two-dimensional coordinates of the landmarks output for each video frame. Then, we used the results of calculating the average value over the entire duration of the mock class.

However, during the mock classes analysis, there were periods when it was hard to detect OpenPose's virtual skeleton or when the virtual skeleton was superimposed on the other person as the audience's body. In this paper, we analyzed these frames by excluding them from the measurement results as noise.

RESULTS AND CONSIDERATIONS

The results of analyzing the acquired data according to the above analysis procedure are shown below.

Subjective Evaluation of Mock Classes

Table 1 shows the results of multiple regression analysis. In this paper, we adopted the forward-backward stepwise method for variable selection. After seven steps of the forward-backward stepwise method, the coefficient of determination of the multiple regression equation was 0.934, and the multiple correlation coefficient was 0.966. From this result, the variables included in the regression equation are "1. The objectives and goals of the class were clearly defined", "4. The class had a good flow", "6. The time allocation was appropriate for the content", and "9. The lecturer's explanation was easy to understand". In particular, in the partial regression coefficient test, the p -value of "9. The lecturer's explanation was easy to understand" was 0.001, and the partial correlation coefficient was 0.671. We found that this

had a significant impact on the overall evaluation. However, in this evaluation sheet, the word “explanation” was used in this evaluation item, and no information was available on details such as whether the teacher’s explanation was verbal, gestural, or a balance between the two.

Table 1. Multiple regression analysis results.

Variable	Partial Regression Coefficient	Standard Error	<i>t</i>	<i>p</i>	Partial Correlation Coefficient
1. Objectives and goals were clearly defined	0.582	0.385	1.512	0.148	0.336
4. The class had a good flow	0.556	0.262	2.119	0.048	0.447
6. The time allocation was appropriate for the content	0.354	0.148	2.390	0.028	0.491
9. The lecturer’s explanation was easy to understand	0.951	0.248	3.837	0.001	0.671
(Constant)	−2.369	1.061	−2.232	0.039	

Motion Analysis of Mock Class Lecturers

Regarding motion analysis, we divided the mock class lecturers into two groups based on the comprehensive evaluation results obtained from the evaluation sheet. One group had a comprehensive evaluation higher than the average value, and the other group had a lower than the average value. Comparing the average values of both groups, we found that the group with a higher average value showed a more considerable value for all parameters: “i. The total amount of movement of both hands”, “ii. The range of motion of both hands”, “iii. The speed of movement of both hands”, and “iv. The total amount of the face direction during the mock class” (Figure 3). However, there were no statistically significant differences in either case by *t*-test. On the other hand, high effect sizes were confirmed for: “i. The total amount of movement of both hands”, and “iv. The total amount of the face direction during the mock class”. Therefore, it is necessary to increase the amount of data and verify the differences in the magnitude of motion more accurately in the next step.

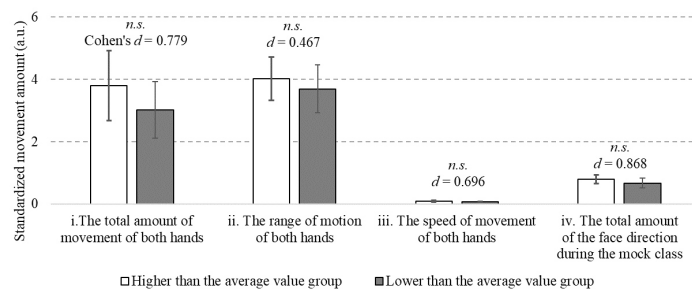


Figure 3: Comparison results of movement amounts among higher and lower groups than the average of a comprehensive evaluation value.

CONCLUSION

In this paper, we conducted a fundamental study on the importance and quantification of how to convey class content, which traditionally relied on qualitative evaluation, in faculty development. Using data from 23 participants' mock classes, we found from their evaluation item that "The lecturer's explanation was easy to understand," which is related to the teacher's delivery method, dramatically influencing the overall evaluation of the mock class. As other motion parameters, frequency of occurrence, the rhythm of change, and the balance and relationship with the voice used simultaneously with gesture explanations are necessary to consider.

In future work, by quantifying the evaluation of delivery methods that have not been quantified, using teachers' motion parameters and combining them with the other quantitative evaluation indicators such as a class design method. Then, it is possible to conduct a comprehensive evaluation of classes objectively. Furthermore, we aim to indicate points that mock class lecturers should improve.

REFERENCES

- Cao, Z., Hidalgo, G., Simon, T., Wei, S., Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 43 Issue 1, pp. 172–186.
- Froyd, J., Fowler, D., Layne, J., Simpson, N. (2005). "Frameworks for faculty development", *Proceedings Frontiers in Education 35th Annual Conference*, Indianapolis, IN, USA. S3E, pp. 23–28.
- Fukuda, M., Huang, H., Nishida, T. (2019). "Detection of Student Teacher's Intention using Multimodal Features in a Virtual Classroom", *Proceedings of the 11th International Conference on Agents and Artificial Intelligence (ICAART 2019)*, Prague, Czech Republic, pp. 170-177.
- Kilic, A. (2010). Learner-Centered Micro Teaching in Teacher Education. *International Journal of Instruction* Vol. 3 No. 1, pp. 77–100.
- Ministry of Education, Culture, Sports, Science and Technology (MEXT). MEXT Website: https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/003/gijiroku/attach/1416153.htm (in Japanese), Accessed 24 September 2023.
- Murray, J. (2020). A Critical Evaluation Towards Micro-Teaching With Suggestions for Future Improvements, *Journal of Foreign Language Education and Research* Vol. 1, pp. 97–109.
- Ozeki, S., Hayashi, T., Fukano, M., Yamazaki, S., Meach, L. A., Sorcinelli, D. M. (2023). Examining the Key Components of Faculty Development to Advance Japanese Higher Education: A Qualitative Study, *International Journal of Institutional Research and Management International Institute of Applied Informatics* Vol. 7 No. 1, pp. 1–16.
- Park, R. (1979). Faculty Development: An Historical Perspective. *POD Quarterly: The Journal of the Professional and Organizational Development Network in Higher Education* Vol. 1 No. 1, pp. 24–32.
- Remesh, A. (2013). Microteaching, an efficient technique for learning effective teaching, *Journal of Research in Medical Sciences*, pp. 158-163.
- UTokyo Faculty Development, The University of Tokyo Future Faculty Program (UTokyo FFP) Programs. UTokyo FFP Website: <https://utokyofd.com/en/ffp/summary/>, Accessed 24 September 2023.