International Journal of Institutional Research and Management
International Institute of Applied Informatics
2021, Vol. 5, No. 1, 65 – 79

Mediation Effect of the Improvement in Teaching Assistant Quality on Students' Evaluations of Their First-Year Experience

Mio Tsubakimoto *

Abstract

In Japan, a first-year experience program is required for new university students with diverse academic abilities and motivations. Proper training and utilization of teaching assistants (TAs) is necessary to ensure the quality of higher education. Since 2015, the University of Tokyo has offered a first-year experience program called the "First-Year Seminar," aimed at promoting problem identifying and solving skills by involving TAs to engage students in the program. However, the causal structure of TAs' effectiveness in class has not been confirmed. This study aims to verify the causal model from the perspective of educational practice using data from students' evaluations of teaching. TA training underwent radical reform in 2017. A multi-group mediation analysis between pre- and post-TA training reform showed that the specific and appropriate support actions of the TA affect students and teachers, thereby influencing students' overall lecture satisfaction. TAs who did not provide enough specific learning support and knowledge had a significantly negative effect on students, faculty, and lecture satisfaction. The educational effects of TA development were found both in self-evaluation by TAs and in the supportive relationships among faculty, students, and TAs. The results of this study demonstrate TAs' work structure toward their goal of supporting classes.

Keywords: first-year experience, learning support, mediation model, student evaluation of teaching, teaching assistant

1 Introduction

With the popularization of higher education in recent years, quality assurance of higher education has become an important issue in education policy. Effective training and utilization of teaching assistants (TAs) is a critical factor of quality assurance. TAs are a kind of academic assistant, described by the National Institution for Academic Degrees and Quality Enhancement of Higher Education [1] as "a worker appointed to assist with the implementation of academic programs." A proposal by Ministry of Education, Culture, Sports, Science and Technology in Japan states that TAs need to develop a sense of responsibility for education and receive postgraduate education that addresses teaching

* The University of Tokyo

methods [2]. The Organization for Economic Co-operation and Development's (OECD) education statistics describe higher education enrollment rates as follows: "If current entry patterns continue, it is estimated that 49% of young adults (excluding international students) will enter tertiary education for the first time before the age of 25, on average across OECD countries" [3]. Additionally, as per the basic school survey in Japan, the percentage of students enrolling in universities in the academic year (AY) of 2018 reached 54.7% [4]. Trow presented a typology of the development of higher education and categorized it into three stages: elite, mass, and universal [5]. According to this classification, higher education in major foreign countries, including Japan, is considered to have reached a stage of universal access through many diverse learners. Therefore, current university students have wide-ranging academic abilities and psychological backgrounds (e.g., motivations for learning). The concept of learning in higher education has also evolved from one-way teaching (teacher-centered) to interactive learning (learner-centered), so that learners can acquire enhanced learning outcomes. This change in perspective led to new learning methods, such as active learning, being adopted. It is likely that first-year university students will experience this form of learning, often in their first-year experience course.

A competent and proficient TA is necessary to implement a successful first-year experience course and active learning. In Japan, as well as the United States, which was the pioneer in offering first-year experience in higher education, the rate of first-year experience courses in universities is now high, based on the universalization of universities. The content of first-year experience courses is very diverse, and includes academic skills, study skills, information literacy, and introduction to academia, among others. Additionally, one important goal of first-year experience is to ensure a smooth transition to university education for new students. To achieve this goal, first-year experience courses need "gimmicks," so to speak, to activate student connections, faculty-student interactions, and access to services on campus. To activate this "trick" effectively, it is imperative to keep class sizes small, introduce active learning, and utilize well-trained TAs.

In an active learning environment, learning is often carried out in a peer or small group learning format, where multiple learners collaborate with each other on tasks [6]. Learners may sometimes be faced with problems that they find difficult to solve on their own, and they may ask others for necessary help. This is called academic help-seeking [7]. Help can be sought from friends in the same group or class, as well as from teachers and TAs. TAs are integral to helping and facilitating student learning activities. The utilization of TAs' help by students is part of an effective learning strategy [8]. TAs can have a positive impact on both the learning process and the outcome of the learner, as learners are more challenged to perform in the presence of their TAs than they are when they are alone [9]. Moreover, their final academic performance is enhanced by the involvement of TAs [10]. TAs can also support the faculty teaching processes. Therefore, the TA is a key person for improving the quality of teaching and learning in university education.

In AY 2015, the University of Tokyo introduced a first-year experience course called "First-Year Seminar" (FYS), which offers all new students a cutting-edge research and learning experience as an introduction to their university education [11]. New students are required to register for one FYS class offered in their specific stream (all new students enter either the natural sciences stream or humanities and social sciences stream), and as a part of this a TA is usually assigned to one or more students. All FYS classes are small (approximately 20–24 students per class). In the natural sciences stream, TA training began in AY 2015 and was conducted annually, although it was not designed considering the specific knowledge and skills required for TAs (e.g., effects of active learning, facilitation technique, and what TAs should and should not do). The training program underwent radical reform in AY 2017 when the aforementioned topics were incorporated into the program. The purpose of this training was to improve the quality of TAs by asking them to acquire appropriate knowledge and skills for learning support, and to develop TAs who can be effective in class.

This study aims to verify the improvement in TA quality using causal modeling from the perspective of learning support using student evaluations of teaching. Two hypotheses were tested: (1) the quality of TAs improved after the training reform, and (2) the quality of TAs affects students' motivation, learning outcomes, support from faculty, and satisfaction with the FYS course. It has not been previously confirmed whether the quality of TAs has improved since the training reform. In this study, "TA quality improvement" is defined as "TA can perform learning support actions that have a positive effect on both students and faculty" for the purpose of both hypotheses. Improvements in TA quality could also have a positive effect on the scores of both teaching by faculty and learning by students. Further, we expected that improved TA quality would have a positive effect on summative class evaluation, mediated by improved faculty and student evaluations.

2 Methods

2.1 FYS Courses

The participants were first-year students enrolled in the natural sciences stream from AY 2015 to 2019 at the University of Tokyo. The total number of first-year natural science students is approximately 1,880 each AY, all of whom register for the FYS course. Students are divided into small classes of approximately 20 students each, and they attended one FYS class on a natural science topic. A total of 100 FYS classes are conducted every year in the natural sciences stream. From all science faculties, 100 (or more) members at the university take charge of the FYS classes based on their specialties; therefore, the themes that students learn vary from class to class. In addition, the manner of utilizing the TAs' assistance by each faculty member and the number of TAs who oversee each class vary.

2.2 TA Training

The content and methods of TA training changed between AY 2015–2017 and AY 2018–2019. As FYS courses started in April 2015, the first TA training took place in February and March 2016, with essentially the same training content in both months. TAs had to attend the training in either February or March, at their convenience, and attendance was voluntary. If TAs attended the training, they were compensated with a time-sensitive honorarium by the university. In the AY 2016 and 2017 training, the content provided to the TAs was the same as that provided to teachers in charge of FYS. The content included classroom design methods and motivation. While this content could be useful for classroom improvement, it was not specific to the work of the TAs.

In contrast, the training content of AY 2018 and 2019 was divided into two categories: one for teachers and the other for TAs. In the TA training, the lecturer first explained the definition of "TA" and the learning effects of working as a TA. The participants (TAs) were then asked to come up with their own ideas for things "to do and not to do by TAs" in a group work style. Finally, they presented their results. The goal of this training was twofold: (1) to let the TAs experience the methods and effects of active learning and (2) to let them identify for themselves the desirable attitudes and behaviors of TAs.

2.3 Participants and Procedure

The data from AY 2015 to 2019 were split into two groups: the data from 2015 to 2017, before improving TA training (henceforth referred to as Pre), and the data from 2018 and 2019, after the improvement (Post). This study will compare these groups to test hypotheses (1) and (2). To make the groups (AY 2015–2017 and AY 2018–2019) as comparable as possible, the exclusion criteria for students were those who (a) had less than 80% class attendance, (b) re-took the FYS course, and (c) had transferred from another university. The faculty in charge of the FYS class asked students to answer a 50-item evaluation questionnaire at the end of the semester (FYS is offered from April to July each year, and in Japan the academic year starts in April). The survey took place in the university classroom where the class was held and was administered separately in each of the 100 classes every year. The response format was mark sensing, and the method used from 2015 to 2017 was unregistered. However, from 2018, the method was changed to a registered Learning Management System. Each year, the following number of students participated in this survey, which was conducted from AY 2015 to 2019: 1,519 (AY 2015), 1,404 (AY 2016), 1,492 (AY 2017), 1,149 (AY 2018), and 1,417 (AY 2019), respectively. The average ratio of valid responses for this survey based on the exclusion criteria was 75.02%. The gender percentage of the participants was not known because it was not obtained as a survey item. However, the percentage of undergraduate female students at Japanese national universities, where the survey was conducted, has been around 20% for the past ten years. According to the World University Rankings [12], this university is ranked 36th in 2021 and third in the Japan University Rankings of 2020. Analysis of the data and publication of the results were permitted by the faculty organization with jurisdiction over this study and by the research ethics committee. The students provided informed consent at the time the survey was conducted.

2.4 Materials and Analysis

This study's assessment measure was a student evaluation of a teaching questionnaire previously developed and used at the university. The questionnaire comprised 50 items evaluating the FYS, and the response formats comprised a combination of Likert scales (e.g., 1: strongly agree to 5: strongly disagree), multiple-choice answers, and open-ended answers. The items consisted of topics such as attendance rates, time management in class, the extent of knowledge acquisition in the class, and learning support acquired from the TAs and the faculty. Of the 50 items, the first 20 were common items, including face sheets also used to evaluate courses other than FYS. The remaining 30 items were especially designed for the purpose of the FYS (e.g., acquisition of scientific skills). Some of these items were deleted or added each year from AY 2015 to AY 2019. For this study's purpose, six items that were used consistently from 2015 to 2019 and assumed to be closely related to the hypotheses were chosen. Of these six items, four were related to common topics (e.g., "You acquired knowledge of the contents in the course," "You were interested in the course," "The teacher answered your questions," and "Overall, you were satisfied with the course") and two were special items ("TA communicated at the level of the students' understanding" and "TA supported group work constructively"). Our data analyses were conducted using IBM SPSS Statistics for Mac Version 25 and Amos for Windows Version 26 (IBM). The dependent variable was overall satisfaction, while the independent variables were TA, teacher, and students' learning (see Items 1–5, Table 3).

3 Results

3.1 Pairwise Analysis

To confirm Hypothesis 1, we conducted a pairwise analysis to determine the differences in the mean of the two items related to TA evaluation (1: the TA communicates at the level of the students' understanding; and 2: the TA supports group work constructively) over the years. The analysis was conducted for each item. The basic statistics for each item and the numbers of valid respondents are presented in Table 1. The results of the pairwise comparisons using Bonferroni corrections indicated significant differences (Post < Pre; detailed *p*-values are shown in Table 2). No significant difference was found between the Post years (2018 and 2019).

Table 1: Basic statistics by AY for the two items

Item	FY	*n*	*M*	*SD*	*SE*	95% CI	
						LL	*UL*
1. TA communicates at the level of the students' understanding	2015	1497	4.16	1.04	0.03	4.11	4.22
	2016	1375	4.24	1.06	0.03	4.18	4.3
	2017	1460	4.24	1.10	0.03	4.18	4.3
	2018	1149	2.17	1.37	0.04	2.09	2.25
	2019	1416	2.20	1.48	0.04	2.12	2.28
2. TA supports group work constructively	2015	1499	4.23	1.02	0.03	4.18	4.29
	2016	1375	4.29	1.05	0.03	4.24	4.35
	2017	1458	4.24	1.13	0.03	4.18	4.29
	2018	1149	2.09	1.35	0.04	2.01	2.17
	2019	1416	2.17	1.48	0.04	2.09	2.25

Note. *n* = number of valid responses; CI = confidence interval; *LL* = lower limit; and *UL* = upper limit. Smaller numbers indicate positive responses.

Table 2: Result of the pairwise analysis for two items

Item	(I) FY	(J) FY	Difference in means (I-J)	*SE*	*p*-value	95% CI	
						LL	*UL*
1.TA communicates at the level of the students' understanding	2015	2016	-0.08	0.05	0.93	-0.20	0.05
		2017	-0.08	0.05	0.80	-0.20	0.05
		2018	1.99*	0.05	0.00	1.86	2.13
		2019	1.96*	0.05	0.00	1.84	2.09
	2016	2015	0.08	0.05	0.93	-0.05	0.20
		2017	0.00	0.05	1.00	-0.13	0.13
		2018	2.07*	0.05	0.00	1.93	2.21
		2019	2.04*	0.05	0.00	1.91	2.17
	2017	2015	0.08	0.05	0.80	-0.05	0.20
		2016	0.00	0.05	1.00	-0.13	0.13
		2018	2.07*	0.05	0.00	1.94	2.21
		2019	2.04*	0.05	0.00	1.91	2.17
	2018	2015	-1.99*	0.05	0.00	-2.13	-1.86
		2016	-2.07*	0.05	0.00	-2.21	-1.93
		2017	-2.07*	0.05	0.00	-2.21	-1.94
		2019	-0.03	0.05	1.00	-0.17	0.10
	2019	2015	-1.96*	0.05	0.00	-2.09	-1.84
		2016	-2.04*	0.05	0.00	-2.17	-1.91
		2017	-2.04*	0.05	0.00	-2.17	-1.91
		2018	0.03	0.05	1.00	-0.10	0.17
2. TA supports group work constructively	2015	2016	-0.06	0.05	1.00	-0.18	0.07
		2017	0.00	0.05	1.00	-0.13	0.12
		2018	2.15*	0.05	0.00	2.01	2.28
		2019	2.07*	0.05	0.00	1.94	2.19
	2016	2015	0.06	0.05	1.00	-0.07	0.18
		2017	0.06	0.05	1.00	-0.07	0.19
		2018	2.20*	0.05	0.00	2.07	2.34
		2019	2.12*	0.05	0.00	1.99	2.25
	2017	2015	0.00	0.05	1.00	-0.12	0.13
		2016	-0.06	0.05	1.00	-0.19	0.07
		2018	2.15*	0.05	0.00	2.01	2.28
		2019	2.07*	0.05	0.00	1.94	2.19
	2018	2015	-2.15*	0.05	0.00	-2.28	-2.01
		2016	-2.20*	0.05	0.00	-2.34	-2.07
		2017	-2.15*	0.05	0.00	-2.28	-2.01
		2019	-0.08	0.05	0.87	-0.22	0.05
	2019	2015	-2.07*	0.05	0.00	-2.19	-1.94
		2016	-2.12*	0.05	0.00	-2.25	-1.99
		2017	-2.07*	0.05	0.00	-2.19	-1.94
		2018	0.08	0.05	0.87	-0.05	0.22

Note. '*' indicates that $p<.05$. Smaller numbers indicate positive responses.

3.2 Factor Analysis

A factor analysis was conducted on the five evaluation items, yielding two factors; next, a Promax rotation was applied to both factors. Table 3 shows the loadings of the items on these two Promax-rotated factors. Factor 1 (F1) is related to TA evaluation and is referred to as "support from TA." Factor 2 (F2) is related to students' self-evaluations and their evaluations of the teacher and is referred to as "support from teacher and learning by students." The eigenvalues of the two factors were 1.94 and 1.80, respectively (the variance was 38.78 and 35.99, respectively). There were no critical issues related to either the value of factor correlation (r = -.04) or Cronbach's alpha (F1: α = .95, F2: α = .67). Although the Cronbach's alpha of F2 showed a somewhat low value, the five items were grouped into the two categories of "TA-related" and "teacher and learner-related," as expected. Further, the interpretability of each factor in the results was not considered to be problematic; therefore, the results were adopted as they were and used for further analysis.

Table 3: Factor analysis of student's evaluation items

Evaluation item	Factor loading	
	1	*2*
Factor 1: Support from TA		
1. TA communicated at the level of the students' understanding.	**.98**	-.00
2. TA supported group work constructively.	**.92**	.00
Factor 2: Support from teacher and learning by students		
3. You acquired a knowledge of the contents in the course.	.02	**.73**
4. You were interested in the course.	.00	**.67**
5. The teacher answered your questions.	-.02	**.55**

Note. Factor loadings above .35 are displayed in bold. There were no reverse-scored items.

3.3 Correlations

To confirm the model assumed by Hypothesis 2, a correlation analysis was conducted for each data group (Pre and Post), based on the results of the factor analysis. Table 4 illustrates the correlations among the variables (F1, F2, and overall satisfaction) using a Pearson's product-moment correlation. F1 was a variable related to learning support by TAs, F2 was a variable related to learning support by teachers and students, and overall satisfaction was a variable related to summative evaluation with the class in the class evaluation questionnaire. A preliminary correlation analysis was conducted to determine any difference in the relationship between these three variables before and after the TA training (Pre and Post). Consequently, there were differences in the positive and negative associations among the variables in Pre, Post, and All years. The results of the "All years" analysis, without separating the years, and the results of the Pre, before the improvement of TA training, both had a significant negative correlation between F1 and F2 (r = -.03, -.33). Additionally, there was no correlation or a significant negative correlation between F1 and overall satisfaction (r = -.01, -.31). In contrast, there was a significant positive correlation between these three variables in Post. These results indicate that the year before and after the TA training was a moderator.

Based on these results, a mediation analysis was conducted. For these three variables, a model was created based on Hypothesis 2 to examine whether there is a difference between the Pre and Post models.

Table 4: Correlation matrix to confirm the multi-group covariance structure

	All years			2015-2017 (Pre)			2018-2019 (Post)		
	F1	*F2*	*OS*	*F1*	*F2*	*OS*	*F1*	*F2*	*OS*
F2	-.03**	-		-.33**	-		.46**	-	-
OS	-.01	.73**	-	-.31**	.72**	-	.45**	.75**	-
M	3.23	1.76	1.77	4.08	1.76	1.80	1.78	1.75	1.78
SD	1.40	0.57	0.86	0.90	0.57	0.87	0.75	0.56	0.84
n	6,138	6,864	6,880	3,856	4,346	4,361	2,282	2,518	2,519

Note. "OS" indicates overall satisfaction. '**' indicates that $p < .01$.

3.4 Mediation Analysis

Multi-group covariance structure mediation analysis was conducted for the two groups (Pre and Post). The mediation model considered that the relation between F1 (TA's support) and overall satisfaction is mediated by F2 (Teacher's support and students' learning). Figures 1(a) and 1(b) display the relationships between F1, F2, and overall satisfaction. Figure 1(a) shows the result of Pre (2015–2017) and 1(b) the result of Post (2018 and 2019). The fit indices of this model were: $\chi^2 = 123.36$, $df = 11$, $p = .00$, $CFI = .99$, $NFI\ Delta1 = .99$, and $RMSEA = .03$. Based on Oshio [13], it was determined that all fit indices were adequate; therefore, this model was adopted.

Each path between pre- and post-training was statistically compared. A significant difference was observed only in the path linking F1 to F2 (Pre = -.39 < Post = .51, $z = 25.54$, $p < .05$). Additionally, both indirect and direct effects were examined using bootstrapping (1,000 trials), which revealed that all paths and effects were significant. The indirect effect was significantly larger than the direct effect (standardized indirect effect: Pre = -.35 [$p = .00$], Post = .47 [$p = .00$]; standardized direct effect: Pre = .04 [$p = .02$], Post = -.06 [$p = .02$]).

The mediation analysis showed that the relationship between F1 and overall satisfaction was partially mediated by F2.

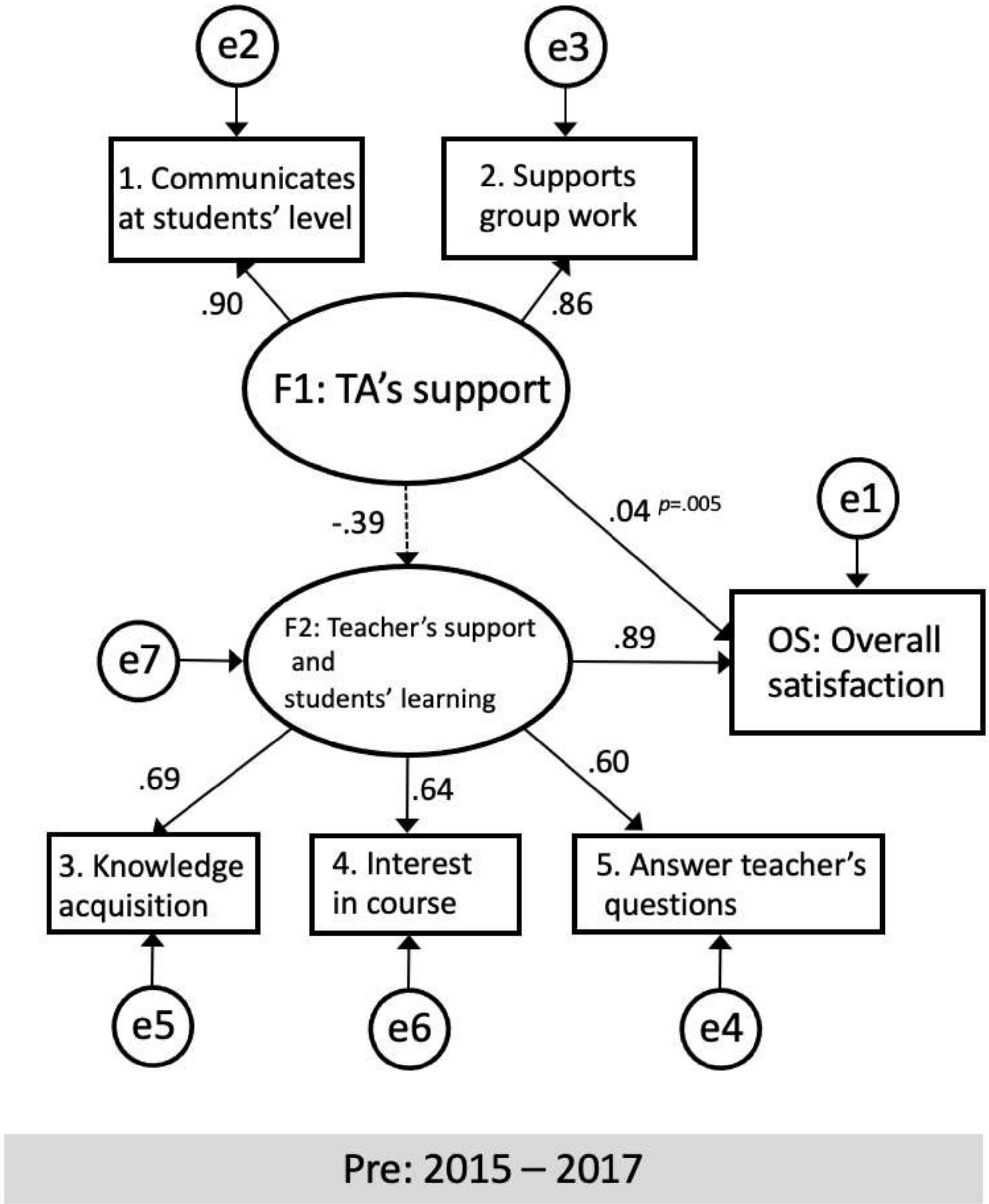

Figure 1(a): Multi-group covariance structure analysis: mediation modeling

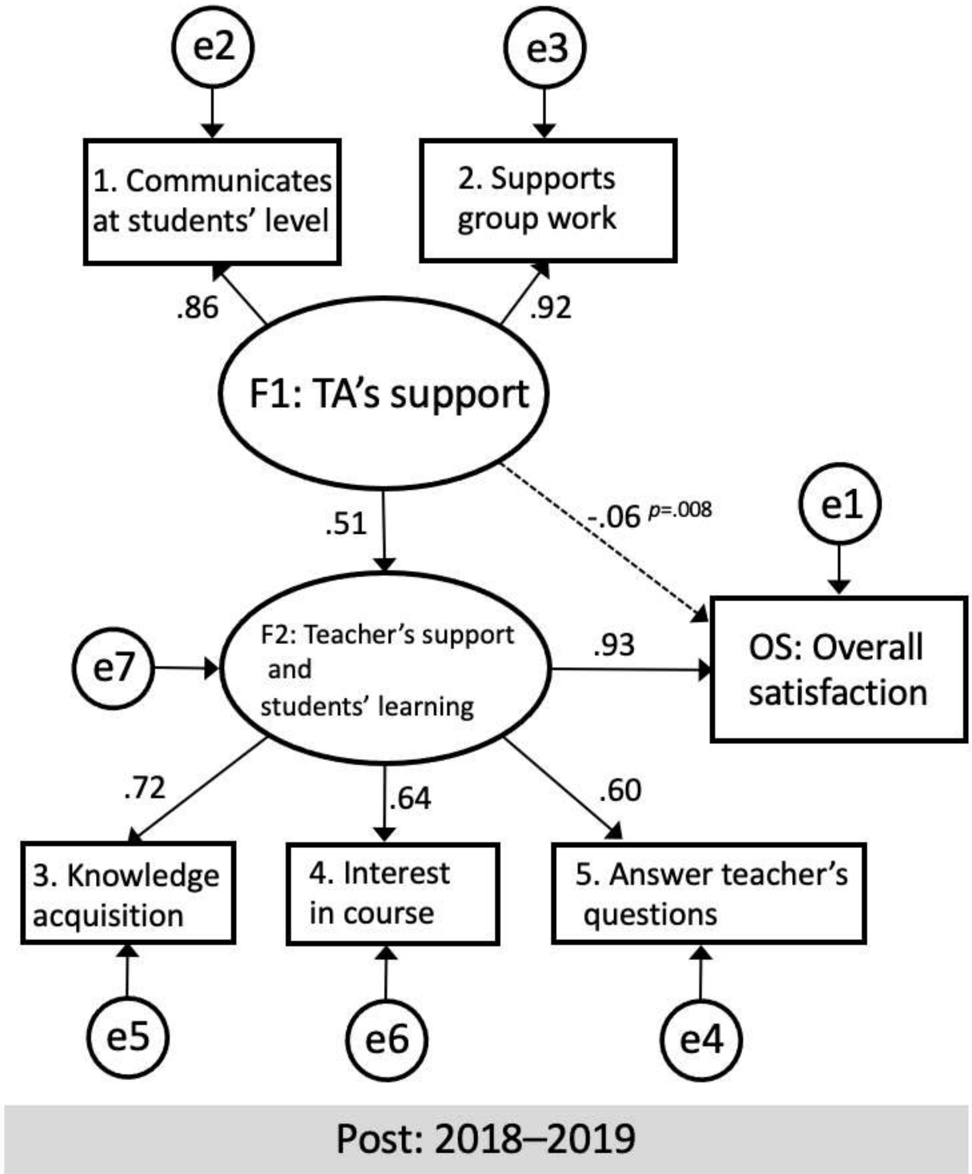

Figure 1(b): Multi-group covariance structure analysis: mediation modeling

4 Discussion

This study focused on a first-year science education course at a national university in Japan. This course has been conducted every year since 2015, and a student's evaluation questionnaire with the same items was taken each year. The items related to the research hypothesis were extracted from the questionnaire data accumulated every year of the study period and used for analysis. Although TA training has been conducted since 2015, the evaluation of TAs in the annual students' evaluation has been relatively low since the first survey.

This study's main purpose is to examine the improvement effect of TA training by using the panel data of the students' evaluations. Our study was designed to evaluate the effects of improved TA quality on improving the quality of teaching and learning. Here, the improvement in teaching refers to improving the quality of teachers' responses to students, while improvement in learning refers to students' acquisition of knowledge and their increased interest in the learning content. The aim of this study was twofold: to test the hypothesis that the quality of TAs improved after the training reform and to test whether the quality of TAs affects students' motivation, learning outcomes, support from faculty, and satisfaction with the FYS course.

This study's most significant finding is that the post-training mediation model had greater educational effectiveness than the pre-training model, thereby confirming the second hypothesis. TA behavior had a weak negative effect on faculty and students before improvements in TA training (Pre) and a positive effect after improvements (Post). There was a significant difference between these negative/positive effects. Further, there was a significant main effect between students' evaluations of TAs, which was related to pre- and post-TA training. This result suggests an improvement in TA training, which supports the first hypothesis. As shown in Table 4, when the relationships between the factors were checked for all years—from AY 2015 to AY 2019—a small correlation was observed between F1 (TA's support) and F2 (teacher's support and students' learning), and between F1 and overall satisfaction. However, when the correlation was reconfirmed by stratifying the data into Pre (AY 2015 to AY 2017) and Post (AY 2018 and AY 2019), a change in the relationship between Pre and Post was observed. This suggests that "year" may have been a confounding factor in our analysis. In an analysis such as the one in our study, which is based on data obtained from a virtually uncontrolled educational setting, it is rational to assume that confounding factors may exist. Whether the analysis policy and results in this study were correct, and whether the confounding factors were handled appropriately, may have to be confirmed by carefully examining the results from a different direction using a different dataset.

The present findings demonstrate that the improvements in TA training were significantly effective in bringing positive changes in TAs' specific support actions for faculty's teaching and students' learning. These findings suggest that an appropriately trained TA demonstrates specific supportive actions in the classroom, which has an important positive effect on both students and faculty, thereby affecting students' overall lecture satisfaction. However, regardless of the quality of the TA, the direct effect of the TA on students' overall satisfaction was less pronounced than the indirect effect. The direct effect being smaller than the indirect effect, and the value of the path coefficient being very small for both Pre and Post, suggests that the direct contribution to the improvement of the overall students' satisfaction with a class is small if there are TAs in the classroom, regardless of the improvement in training. In a real classroom, it may be important for TAs to be actively involved with teachers and students.

TAs are factors that affect overall satisfaction only when they are involved with the faculty and students. Therefore, it may be important to maintain the appropriate quality and quantity of teaching and TA learning support behaviors.

Previous studies have demonstrated that the presence of TAs improved learning achievements as well as students' attitudes [8–10]. However, in this study, the result of mediation modeling (Pre) was different from previous studies (-.39 negative effect from TAs to faculty and students). This result is very interesting. To obtain a positive educational effect, as in previous studies, there must have been an assumption that the TAs should be well trained. To make effective use of TAs, it is important to train them to understand what they can and should do to improve education and be willing to implement effective learning support actions in their classrooms.

Previous studies have demonstrated that both the students being taught and the TAs experience learning effects. For example, Comfort and McMahon found that students who taught others exhibited significantly higher academic performance than those who did not [14]. Their study was conducted with undergraduate students; therefore, its applicability to this study, where most TAs are graduate students, might be limited. However, the effect of "Learning by Teaching" may be widely applicable. For TAs to experience this learning effect to the fullest, it is important to ensure the quality of TA training. To be more precise, it is necessary to provide training on appropriate and specific learning support behaviors among TAs. This is because this educational knowledge could help TAs understand the appropriateness of their own actions as TAs and their own learning process as learners. This may promote their own metacognition and improve the quality of both learning and learning support.

5 Limitations and Further Viewpoints

This study has some limitations. First, serious issues have been pointed out regarding the validity of the scores of the student evaluations of teaching, which were used to test the hypotheses. Such evaluations can be highly skewed and are not a reliable measure of educational effectiveness [15]. Therefore, the scores may not be well related to students' learning outcomes. Student evaluations were used in this study not because the author believes that student evaluations can measure student performance. Student performance data obtained continuously over many years and recorded on a higher-order scale—which is necessary to achieve the purpose of this study more clearly—did not exist in the class that was the subject of this study. In addition, each of the 100 courses for this class has different content, and the content and number of exercises were not constant. The class that was the subject of this study was not a class designed for research, but a "real world" situation; therefore, there were significant limitations to the author's research intervention. Considering the above, this study attempted to test the hypothesis using the results of a student survey, which can be obtained relatively effortlessly and over time in the field of education.

However, the two questions related to TA evaluation used in this study are related to the extent to which the TA has performed specific learning support actions. Therefore, the scores obtained reflect the facts relatively well. Nevertheless, a variety of educational data sets need to be used to validate this study's results. For example, to measure student performance more accurately, cloud collaboration tools such as Google Docs could be used to visualize the communication and thinking processes of students as much as possible, after which they can be numerically quantified using rubrics. With the recent growth of online classes in higher education due to the COVID-19 pandemic, this method might be relatively easy to implement. Further, clearer measurements of TAs' performances may also be necessary. It is therefore necessary to measure the learning support behavior of TAs. Tsubakimoto [16] conducted a qualitative survey of TAs for this purpose and developed a prototype rubric.

Second, the final grade of the FYS students in natural sciences is expressed simply in terms of pass or fail, and more than 90% of the students pass each year. These highly biased binary data are very difficult to use effectively for analysis.

Third, there are three biases in the data collected. First, although it was mandatory to answer the class evaluation questionnaire, some students did not provide answers. Therefore, the author could use only the data of the students who responded to the questionnaire and could not obtain a complete enumeration. Additionally, because of the nature of this questionnaire, random sampling was not possible. Second, there is a bias in the ratio of male and female students who responded to the questionnaire. The ratio of male to female students at the university where the survey was conducted was skewed from the start. This may have reflected in the data. However, as the data were collected anonymously, the degree to which they reflect this bias is unknown. Finally, the data presented here are limited to those from first-year education classes. The students who responded to this survey were first-year undergraduates and had little experience with TAs. It is important to remember that this lack of experience may have reduced the validity of the evaluation for TAs.

Fourth, there is a lack of a multifaceted evaluation. In this study, the author analyzed only the data from the questionnaire for a class evaluation by the students. However, how much can we claim has the quality of TAs improved, based on these data from only students? In the future, it will be necessary to collect data on the quality of TAs from both TAs and faculty members.

The fifth problem is that the results of the improvement in the quality of TAs are not pedagogically concrete. What kind of learning TA support behaviors were effective and how were learning situations not the subject of this data analysis? Clarification of effective learning behaviors and the situations in which they were effective would be useful information for refining the TA training content. It will be necessary to conduct a qualitative survey of TAs to collect the necessary data in the future.

The last problem is that this study consisted of a case study of only one university. Readers should take into account the fact that the findings of this study are limited in this sense as well. In the future, it will be necessary to clarify the extent to which the results can be generalized by comparing the results with those of other universities that use TAs in first-year education or with other classes within the same university.

The results obtained in this study may change in the future, if, for some reason, the annual TA training cannot be conducted satisfactorily. Only less than half of the usual number of TAs could be trained in AY 2020, due to the impact of the COVID-19 pandemic. This impact needs to be confirmed by ongoing analyses. From the perspective of educational practice, it will be important to continuously update the training programs, so that TAs can be trained to take effective learning support actions based on changing classroom situations such as online, face-to-face, and highly flexible teaching situations.

Acknowledgment

The author gratefully acknowledges the collaboration provided to develop the TA training program by the members in the divisions of First-Year Education, Active Learning and Teaching, and Center for Research and Development of Higher Education at the University of Tokyo. The author would like to thank the students for consenting for the use of these data within this study. This study was supported by JSPS KAKENHI Grant Number JP17K01086 and 19H01710.

References

[1] National Institution for Academic Degrees and Quality Enhancement of Higher Education. *Glossary of Quality Assurance in Japanese Higher Education*, 4th ed. (2016). Accessed: Apr. 19, 2021. [Online]. Available: https://www.niad.ac.jp/n_kokusai/publish/no17_glossary_4th_edition.pdf

[2] Ministry of Education, Culture, Sports, Science and Technology. "Major Recommendations from the University Council Reports on Teaching Assistants (TAs)." Ministry of Education, Culture, Sports, Science and Technology. https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/003/gijiroku/07011713/001/002.htm (in Japanese) (Accessed Apr. 19, 2021).

[3] OECD. "indicator B4. Who is expected to enter tertiary education?." Education at a Glance 2020: OECD Indicators. https://www.oecd-ilibrary.org/sites/69096873-en/1/3/3/4/index.html?itemId=/content/publication/69096873en&_csp_=b68ed930151b9d0f354aa48ee27178db&itemIGO=oecd&itemContentType=book#boxsection-d1e15412 (Accessed Apr. 19, 2021).

[4] Ministry of Education, Culture, Sports, Science and Technology. "A summary of basic school survey." http://warp.ndl.go.jp/info:ndljp/pid/11293659/www.mext.go.jp/component/b_menu/other/__icsFiles/afieldfile/2018/12/25/1407449_2.pdf (in Japanese) (Accessed Apr. 19, 2021).

[5] M. Trow, "The Expansion and Transformation of Higher Education," The Notion of Modern Educational Sociology, vol. 18, no. 1, 1972, pp. 61-84.

[6] R. Ikeda, and Y. Tateoka, *Introduction to Peer Learning*, Tokyo, Japan: Hituzi Shobo (in Japanese), 2007. ISBN: 978-4894762886.

[7] M. Nakaya, "Children's Social Responsibility Goal, Learning Behavior, Academic Achievement in Classrooms," Journal of Educational Psychology; doi: 10.5926/jjep1953.46.3_291. (in Japanese)

[8] S. Nelson-Le Gall, "Help-seeking behavior in learning," Review of Educational Research, vol. 12, 1985, pp. 55-90.

[9] Goldschmid, B., and Goldschmid, M., "Peer Teaching in Higher Education: A Review," Higher Education, vol. 5, no. 1, 1976, pp. 9-33.

[10] A. M. Ryan, and H. Shin, "Help-seeking tendencies during early adolescence: An examination of motivational correlates and consequences for achievement," Learning and Instruction, vol. 21, 2011, pp. 247-256.

[11] The University of Tokyo. "KOMEX Division of First-Year Education." The University of Tokyo, Komaba. https://www.c.u- tokyo.ac.jp/eng_site/info/research/organization/komex/fresh-year/ (Accessed Apr. 19, 2021).

[12] THE World Universities Insights. "World University Rankings 2021." Times Higher Education. https://www.timeshighereducation.com/world-university-rankings/2021/world-ranking#!/page/0/length/25/sort_by/rank/sort_order/asc/cols/stats (Accessed Apr. 19, 2021)

[13] A. Oshio, *Psychological and Survey Data Analysis by SPSS and Amos*, 3rd ed. Tokyo, Japan: Tokyo Tosho (in Japanese), 2018. ISBN: 978-4489022913.

[14] P. Comfort, and J. J. McMahon, "The effect of peer tutoring on academic achievement," Journal of Applied Research in Higher Education, vol. 6, 2014, pp.168-175.

[15] B. Uttl, C. A. White, and D. W. Gonzalez, "Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related," Studies in Educational Evaluation, vol. 54, 2017, pp. 22-42.

[16] M. Tsubakimoto, "Prototype of a Rubric for Classifying Teaching Assistants' Learning Support Behaviors by Social Channels," unpublished.