

THE ANALYSIS ON TEACHING SKILLS AND LEARNING EFFECTIVENESS OF INTERNSHIP STUDENTS

Abdul Rachman Syam Tuasikal *, Setiyo Hartoto, Bayu Budi Prakoso, Dwi Cahyo Kartiko, Agus Hariyanto,

Universitas Negeri Surabaya, Surabaya, Indonesia

*e-mail: rachmantuasikal@unesa.ac.id

Abstract: The effectiveness of implementing *Pengenalan Lapangan Persekolahan*, new program of introduction to teaching internship, for PE students, is still questionable, starting from its impact on teaching skills, learning effectiveness, and students' teaching load. Meanwhile, this point is significant in *Kurikulum Merdeka Belajar Kampus Merdeka* (MBKM). For this reason, this analytic study will criticize MBKM by providing evidence of the effectiveness of existing internships in the form of teaching skills, learning effectiveness, and students' teaching load. A total of 32 students were selected using consecutive sampling. Teaching skills was measured by observing 12 indicators and learning effectiveness was measured using *Formative Class Evaluation* (FCE) instrument. Anova Repeated Measure was used to analyse the data. The results show that the new program of introduction to teaching internship is not practical, proved by the progressive improvement of teaching skills followed by the ineffectiveness of the learning process. Students feel they have an overload teaching schedule, so that self-development activities are minimal, even though the tutor teachers have guided optimally. It can be concluded that the new program of introduction to teaching internship is less effective because of the overload teaching schedule in such limited time. It is then expected that the longer internship duration would enrich students' experiences under the supervision and more comprehensive teaching load regulation.

Keywords: Kurikulum Merdeka Belajar Kampus Merdeka, teaching internship, physical education teacher candidate, learning effectiveness, and teaching skills.

ANALISIS TERHADAP KETERAMPILAN MENGAJAR DAN EFEKTIVITAS PEMBELAJARAN PADA MAHASISWA MAGANG

Abstrak: Keefektifan pelaksanaan pengenalan lapangan persekolahan (PLP/magang) mahasiswa calon guru di sekolah masih dipertanyakan mulai dari dampaknya terhadap keterampilan mengajar, efektivitas pembelajaran, dan beban mengajar mahasiswa. Padahal keterangan tersebut sangat penting saat kurikulum Merdeka Belajar dan Kampus Merdeka (MBKM) akan memberlakukan program magang di sekolah dilakukan selama satu semester penuh. Untuk itu, penelitian ini mengajukan kritik terhadap MBKM dengan memberikan bukti keefektifan magang yang sudah pernah ada berupa keterampilan mengajar, efektivitas pembelajaran, dan beban mengajar yang didapatkan mahasiswa. Sebanyak 32 mahasiswa terpilih berdasarkan *consecutive sampling*. Variabel keterampilan mengajar diukur dengan cara observasi terhadap 12 indikator dan efektivitas diukur menggunakan angket *Formative Class Evaluation* (FCE). Pengukuran dilakukan sebanyak 3 kali (minggu awal, tengah, dan akhir). Analisis data menggunakan Anova *Repeated Measurement*. Hasil menunjukkan bahwa magang terlaksana tidak efektif dibuktikan dengan keterampilan mengajar meningkat secara progresif sedangkan efektivitas pembelajaran tidak. Mahasiswa merasa memiliki beban mengajar terlalu tinggi sehingga kegiatan pengembangan diri sangat terbatas walaupun guru pamong telah optimal melakukan pembimbingan. Dapat disimpulkan bahwa pelaksanaan magang kurang efektif karena waktunya terbatas tetapi beban mengajar tinggi. MBKM memberikan harapan semakin banyaknya pengalaman yang didapatkan mahasiswa tetapi dengan syarat pengawasan dan pengaturan beban mengajar yang lebih komprehensif.

Kata Kunci: Merdeka Belajar Kampus Merdeka, program magang, mahasiswa calon guru pendidikan Jasmani Olahraga dan Kesehatan, efektivitas pembelajaran, keterampilan mengajar

INTRODUCTION

Teachers are the government policies' implementation holders in the field of education. However, in the last ten years, the government is

still facing teacher competence, teacher recruitment, teacher distribution, and teacher welfare (Ganefri, 2016). When compared to what happened in America, there was also a shortage of

teachers. Still, they immediately thought of a solution. They ran a program in the form of community teacher involvement in preparing students who were doing internships to be more familiar with the routine and quality development of physical education classes in schools (Bert & Richards, 2018). The teaching internship program is a flagship to increase the number of new teachers coming from universities.

Teaching internship is one of the most important programs for physical education students who must directly involve them in the teaching field. Their teaching practice is not sufficient by just having microteaching. The program is one of the leading educational study programs in providing learning experiences for students in managing classes directly at school (Nagro et al., 2017). Besides, this teaching internship program is used to test the suitability of prospective higher education outputs with the demands of employment. As has been developed in Europe, the higher education effectiveness assessment system is carried out in three phases, namely: (1) theoretical competency evaluation based on the applicable curriculum policy, (2) evaluation of the quality of the learning process and the types of competencies that are formed according to field needs; and (3) accurate evaluation in the realm of learning (Bergsmann et al., 2015). It seems that the second phase is the most suitable phase to be conducted in the form of a teaching internship.

The processes of learning by doing and evaluation are very synonymous in the implementation of teaching internships. Implementing the teaching internship will never be far from the process of accustoming students to school culture and evaluating aspects related to learning. For this reason, research related to student's internship is often carried out to evaluate the process of implementing internships in terms of student-teacher performance. Research in the last decade was conducted to evaluate the implementation of internships for physical education students at the same college. The study was conducted to evaluate students in educational and noneducational departments, which shows no difference in teaching and learning effectiveness based on teachers' backgrounds. However, the quality of the learning process carried out by the physical education students during the internship is in the excellent category have various abilities in mastering curriculum, sports, and science in teaching (Khory et al., 2018). Then, it can be concluded that advanced monitoring on the quality of internships has not been concerned yet, especially in

terms of the students' quality and the learning effectiveness.

The tutors' monitoring is essential in responding to the government's new policy in the MBKM. There are four main policies in MBKM for higher education, namely: (1) opening of new study programs; (2) higher education accreditation system; (3) legal entity state universities; and (4) the right to study three semesters outside the study program (Tohir, 2020). The introduction to teaching internship for physical education students will be held for one entire semester with the fourth policy principle. This policy is expected to impact the students' depth of experience and knowledge from the school environment in real terms. Through this policy, the activity will drastically change the internship duration, from 2 - 3 months or even more to an entire semester. The results of this research are expected to provide the basis for policy development at the tertiary level so that they are better prepared to develop the implementation of the introduction to teaching internship MBKM model.

This study, therefore, aims to present empirical evidence of the effectiveness of the teaching internship program carried out by universities, which can then be used to criticize the MBKM-based internship program, which is completely different from the previous ones. Present study is important because it provides the progress of the quality of PE teacher candidates when carrying out the internships. The analysis that was done in this study was also more comprehensive than researches that have been done previously. This can be seen from the number of measurements that are carried out at the beginning, middle, and end of the internship and also in terms of the measured variables, such as teaching skills and learning effectiveness, which both of them can be used as an indicator of the effectiveness of the internship implementation.

METHODS

This analytic study involved 32 students of 7th-semester physical education students taking internships at State University of Surabaya's partner schools. The sample was selected using the consecutive sampling method from a population of 170 students who carried out internships in Surabaya. Assessment of the quality of internships is carried out using observation by the tutor teachers and students' opinions.

Data collection was carried out using two instruments, namely: (1) observation sheets of 12

teaching skills; and (2) a questionnaire for students in the form of formative class evaluation (FCE). The observation sheet of 12 teaching skills used in the study had fulfilled the content validity based on the results of previous research coming from ministry documents regarding the curriculum that has been reviewed from the experts (Suroto & Khory, 2015). It uses a rating scale of 1-5 for each indicator, and the final value is determined using the proportion formula so that the value obtained is 0-1, where higher number represents higher students' teaching skills. FCE is an assessment questionnaire of the physical education learning process that utilizes students' opinions immediately after completing the learning process. This questionnaire is believed to have sufficient validity and reliability based on previous studies (Hasegawa et al., 1995; Suroto & Takahashi, 2005). In Indonesia, the FCE instrument had been tested and the item validity results was found to be 0.6 – 0.83, which fall in the “moderate” to “excellent” criteria (Wijaya & Astono, 2006). This assessment consists of 9 question items with the possible answers of yes (3), do not know (2), and no (1). The final score is from the average of students' entries in a class, and the final score is in the range of 1-3.

Data collection was carried out by mentoring students by a team of researchers who have been trained in research and the use of research instruments. They are sports education department lecturers who are assigned to visit partner schools directly. Assistance and tutorial are provided for students regarding the procedures for using research instruments for data tabulation. In addition, assistance is provided to ensure the variable measurement is carried out according to a predetermined schedule. Data collection was carried out for eight weeks during the internship period. The first measurement was carried out at weeks 1 and 2, the second measurement was carried out at weeks 4 and 5, and the third measurement was carried out at weeks 7 and 8. Data analysis used descriptive statistics to determine data distribution based on skewness ($-1 \leq x \leq 1$) (Meyers et al., 2013). Anova repeated measurement was used to determine the progress of increasing teaching skills and the effectiveness of learning by selecting tests that match the needs of the analysis. All tests will use a significance level of 0.05.

FINDING AND DISCUSSION

Finding

The teaching skills variable used a scale of 0-1, the effectiveness of learning — Formative Class Evaluation (FCE) used a scale of 1-3, and the competency variable used a scale of 0-1. The results of the descriptive statistical analysis of the teaching skills variable were presented in Table 1. The average of first measurement was 0.85 ± 0.095 , measurement 2 was 0.87 ± 0.087 , measurement 3 was 0.88 ± 0.092 , and the average of all measurements was 0.87 ± 0.087 . The results of the descriptive statistical analysis of the learning effectiveness variable or FCE showed that the mean value of measurement 1, 2, and 3 were 2.72 ± 0.134 , 2.74 ± 0.124 , and 2.75 ± 0.109 , respectively, while the average of all measurements was 2.74 ± 0.097 .

The skewness value for all existing variables is at a value of -0.228 to -0.934, this indicates that the shape of the data distribution of all variables is symmetrical ($-1 \leq x \leq 1$) or it can be concluded that the form of data distribution is normal (Meyers et al., 2013). Normality test using Kolmogorov Smirnov also showed that all variables are normal ($p > 0.05$). Based on these results, the analysis can be continued using parametric statistics for the teaching skills variable and the effectiveness of learning based on the FCE questionnaire using Repeated Measurement Anova. The following Table 2 presented Mauchly's Test of Sphericity used to determine the homogeneity variance of the data.

Mauchly's Test of Sphericity revealed that FCE variable has met the homogeneity requirements, but the TS variable did not ($p < 0.05$). For this reason, for the TS variable, Greenhouse-Geisser Tests will be used as a benchmark for determining conclusions while the FCE variable uses sphericity. Table 3 contains the challenges of the univariate test results for the TS and FCE variables. Based on Table 3, it can be explained that one-way repeated measurement analysis of variance (ANOVA) was carried out to test the null hypothesis in the form of no difference in the mean values of TS and FCE in the three groups of time intervals measurement. The ANOVA result for the TS variable is $F = 6.658$, $p < 0.01$, $\eta^2 = 0.117$, indicating that the null hypothesis is rejected. When the alternative hypothesis applies, there is a difference in the average TS value in the three groups of time intervals. In contrast to the FCE, the result is $F = 1.173$, $p > 0.05$, $\eta^2 = 0.316$, indicating that the null hypothesis is accepted; that is, there is no difference in the mean score of FCE in the three groups of time intervals. Based on the results of this analysis, the Pairwise

Comparisons Teaching Skills (TS) analysis can be analysed further, which is illustrated in Figure 1.

The Pairwise Comparisons chosen are the mean value comparisons for the non-homo-geneous variants. It can be seen in Figure 1 that there was an increase in TS from measurement 1 to

measurement 2. Based on statistical tests, the increase occurred significantly ($\Delta = 0.021$, $p < 0.05$). Likewise, there was an increase from the 2 to 3, but it was not significant ($\Delta = 0.008$, $p > 0.05$). When compared to measurements 1 and 3, there was a significant increase ($\Delta = 0.029$, $p < 0.05$).

Table 1. Descriptive Statistics of The Teaching Skills, FCE, and Competency Variables

Variabel	N	Mean	Std. Deviation	Skewness	Sig.*	Scale
TS1	32	0.85	0.095	-0.893	0.120	0-1
TS2	32	0.87	0.087	-0.424	0.200	0-1
TS3	32	0.88	0.092	-0.331	0.075	0-1
Total_TS	32	0.87	0.087	-0.606	0.139	0-1
FCE1	32	2.72	0.134	-0.911	0.110	1-3
FCE2	32	2.74	0.124	-0.228	0.200	1-3
FCE3	32	2.75	0.109	-0.934	0.165	1-3
Total_FCE	32	2.74	0.097	-0.489	0.063	1-3

*) Normality test using Kolmogorov Smirnov ($\alpha = 0.05$)

Table 2. Mauchly's Test of Sphericity

Variable	Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
Teaching skills (TS)	Time	.588	15.948	2	.000
Formative class evaluation (FCE)	Time	.906	2.955	2	.914

Table 3. Univariate Tests

Variable	Type III Sum of Square	df	Mean Square	F	Sig.	η^2
Teaching skills (TS)*	0.14	1.416	0.010	6.658	0.007	0.117
Formative class evaluation (FCE)**	0.20	1.828	0.011	1.173	0.313	0.036

Note: *= Greenhouse-Geisser Tests; **= Sphericity assumed

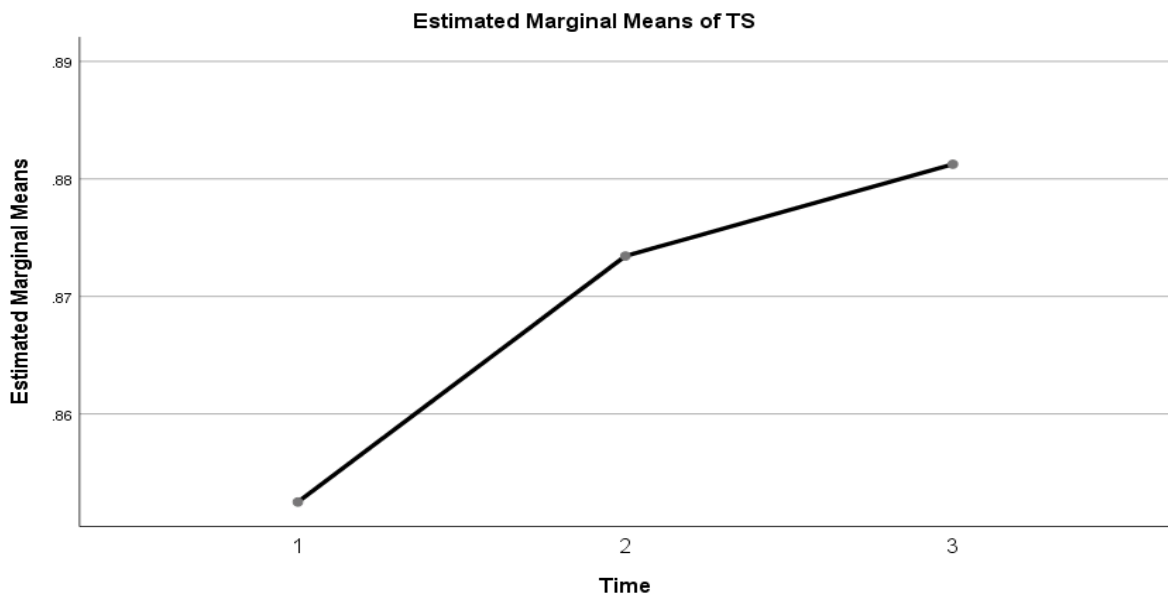


Figure 1. Pairwise Comparisons Teaching Skills (TS)

Discussion

The null hypothesis on the analysis results of increasing teaching skills (TS) was successfully rejected. So, it can be written that there was

a significant increase in the time interval measurement during the teaching internship period. This finding is supported by other research that states that internships can significantly improve student teaching skills (Nagro et al., 2017). The results of this study are not merely about showing the success of internships in improving teaching skills, but rather on the efforts of universities to give the impression of students to realize that internships are carried out to provide teaching experience to students. For this reason, in this study, students make more use of videotapes or analysis of learning videos independently to dig the students' reflective process.

As if following this research, this research also wants to give a reflective impression to the students. However, differently, the reflection is carried out with the tutor's help through the observation process of providing an assessment of teaching skills. The results of TS data analysis obtained from the tutor teacher's observations showed a significant increase occurred from 1 to 2 and 1 to 3; unfortunately, the measurement 2 to 3 was not significant. Although different, this study believes that tutor teachers have each objective grade of value based on existing indicators. So that the tutor teacher gives the value as is according to the performance shown by the students. However, we do not rule out that subjective scores will emerge from the tutors. However, we also believe that statistical analysis provides a fairer choice in addressing the existence of subjective or objective nature of the assessment itself by considering the tolerance for error in getting the conclusions closest to the truth (Gelman & Hennig, 2017). For this reason, the results of statistical analysis of the assessment given by the tutor teacher can still be justified.

The learning effectiveness variable through the FCE questionnaire showed an insignificant increase in contrast to the teaching skills. The first possibility is that the data collection was too short. The time interval may be relatively short so that students cannot give a different assessment when the assessment takes place. These findings also reinforce the opinion that it is necessary to have sufficient time intervals for the research samples to respond differently in each measurement (Guo et al., 2013). Repetition of this kind of response is natural because the instruments used in each measurement are the same.

These technical problems also seem to cause the absence of a relationship between the effectiveness of learning and teaching skills and competence. It should be based on the results of

other researches. It can be proven that high-quality teachers can only create quality learning (Suroto, 2005). The higher the teacher's competence, the higher the quality of learning carried out. So it is very natural that the government is very concerned about the quality of education in Indonesia when the national average teacher competency is 53.02 (below the national target of 55.00) (Maulipaksi, 2016).

Investigating the quality of internships based on teaching skills and the effectiveness of learning is an effort to provide students with teaching experience in terms of self-reflection skills. In fact, this reflective process with the tutulary teacher aims to make students and tutors work together to improve the quality of learning while improving the teaching skills of prospective teacher students. In addition, in this way, it allows students and tutor teachers to review the existing rules in the apprenticeship manual and implement them together. Based on the research findings, some students admitted that civil service teachers did not follow the applicable regulations in the apprenticeship process. The comments given by the students were as follows: "*I hope to have a tutor teacher who follows the order according to the guidebook given by the field supervisor.*" Moreover, violations of rules can be encountered, and a severe problem of carrying out many teaching loads might be possible. The admission of students during the internship is as follows: "*I am happy to get a lot of experience in the internship, we are happy to study with students every day for at least 24 hours of subjects a week.*" According to government regulations, the duration of active learning at school in a week is 40 hours. Teaching for 24 hours means that students must teach 8 classes a week. Due to schools' limited indoor sports halls, physical education classes are mostly scheduled for the first 6 hours in the morning. It indicates that they actively teach for four days, which refers to one left active day for self-development.

This condition is considered a challenging condition if carried out continuously because students will have very few opportunities to develop themselves. Other possibilities are the too high teaching load obtained by students in addition to the technical possibilities of data collection (subjectivity and short measurement time intervals). It might cause limited student self-development. This apprenticeship program should be carried out by maximizing collaboration between student-teacher candidates and tutor teachers. There is an accurate transfer of knowledge and

experience in getting self-development between the two (Hill & Brodin, 2004). With a high teaching load, the interned students may experience difficulties in teaching, especially if it is added to the reflective process given by the tutor teacher, which requires them to have higher competence in responding to improving themselves according to the input given by the tutor teacher. These aspects prove the failure to implement student-teacher internships carried out with a low level of collaboration between students and civil service teachers (Marcon et al., 2012).

Furthermore, by reading students' comments in appreciating the opportunities provided by the school, we can conclude that students are happy with a high teaching load. However, problems will arise if there are more interned students than the tutor teachers so that clinical monitoring of the learning process carried out by teachers is not appropriately evaluated. This condition of delay in clinical observation is a classic problem that has occurred for a very long time (Rikard & Knight, 1997). It is considered a severe problem that has not been resolved yet.

Implementing an internship for eight weeks is considered very short in providing students with experiences about the school environment and real teaching habits. It can also not be used to measure the effectiveness of learning and the quality of the teaching process carried out by students. The real essence of the quality of physical education learning takes at least 20 meetings (Ward, 2013). In addition, other research results indicate that the time required for students to gain understanding and stability in starting classroom teaching activities in natural conditions in the field is 15 weeks (Barney & Pleban, 2006). The presence of MBKM by mandating universities to give students the right to study at the apprenticeship for one semester raises new optimistic hopes for the quality of prospective teacher students. Colleges that focus on developing physical education students should have a more challenging and better effort to have teacher abilities under national standards and comply with government regulations and develop them (Ward, 2016). When implementation of government policies is still not optimal, then the universities could be considered failed agencies to prepare students to become effective teachers (Şen, 2010).

In order to avoid this failure, the moment of MBKM should be used by universities to formulate strategies so that the objectives of apprenticeship can be achieved. The next MBKM-based internship projection must pay attention to five

important things, namely: (1) readiness of an integrated assessment instrument; (2) expertise of field supervisors; (3) the competence of civil service teachers; (4) quality of partner schools; and (5) internship incentives.

The first is the integrated assessment instrument readiness. Unlike the previous internships, the MBKM-based internship will have a higher course unit load of 20-40 course units. In its implementation, it will integrate several courses that support the implementation of internships at schools. The existence of an assessment instrument indicates that the purpose of the internship has been clearly defined. The experience and knowledge obtained are well structured so that the instrument will be useful as a checklist during the internship. The arrangement of instruments should be mutually integrated for all subjects included in the list of course unit conversions in the apprenticeship. Subjects that are included in student apprenticeships include at least the student's performance in managing four general areas of physical education teaching habits, namely: (1) understanding the curriculum; (2) learning environment; (3) teaching strategies; and (4) assessment (Barney et al., 2012). An understanding of the curriculum should require students to compile a physical education learning curriculum in schools. Understanding the learning environment requires apprentice students to adjust the various lesson plans included in the curriculum with school conditions, including the available sports facilities and infrastructure. The teaching strategy requires students to have classroom management skills, prepare for classes, and make students learn (Muhtar & Dallyono, 2020). The assessment relates to the apprentice student's ability to assess student learning progress to improve planning to implementation of learning.

The second is the expertise of field supervisors. One field supervisor carries out assistance to groups of interned students, whereas students are from different departments in one internship group. Since the supervisor will be very unlikely to master all the subjects in the group, it is suggested that each student from a different department might have a particular supervisor. It could be done by having collaboration between universities and schools as mutual monitoring of student internship activities.

The third is the tutor teachers' competence. The most apparent hope in implementing the MBKM-based internship is the assistance of civil service teachers to students. Tutor teachers should have a role as the primary supervisor in

assessment and mentors for students in learning to develop themselves, competencies, and professionalism (Fransson & K. McMahan, 2013). To play the role of supervision and mentoring, the tutors should have high competence to improve students' quality of teaching and learning activities (Marli, 2012). It does not mean that this research suggests collaboration between low competence tutor teachers, but the availability of great tutor teachers is very much needed in the implementation of internships.

Fourth is the quality of partner schools. The hope of implementing the MBKM-based internship is to get high school quality. School quality, in general, can be based on the quality of the principal's leadership (Clifford et al., 2012). The principal's background has a role in determining teacher job satisfaction which has implications for teacher quality (Shen et al., 2012). Even the principal is believed to be an important factor in reducing the student achievement gap in a school (Dhuey & Smith, 2014). It means that by getting an internship in the form of a highly qualified school, students will get ideal learning demands. Besides, they would also get a picture of ideal school conditions, classrooms, and learning processes before facing various problems.

Last is the internship incentives. As being teachers are very hard, they should be respected worthfully. Other research results show that the incentive system positively impacts teachers' performance (Imberman & Lovenheim, 2015). By increasing the time for conducting MBKM-based internships, the cost of implementing the internships should increasingly burden the university finances. For state universities that get direct funding from the government, it should not be a problem. However, for private universities, it will be different. A possible step that could be taken is to collaborate with local governments in providing incentives through a credit point mechanism. It can be proposed in the incentive system for local government or higher education institutions to provide non-profit awards that can be converted into tutor tutoring careers. By accompanying apprentice students, tutors can carry out two tasks at once, namely the task of self-development and the task of participating in preparing students to become great candidates (Faikhamta & Clarke, 2018).

CONCLUSION

Introduction to teaching internships is an important program in introducing the school environment to fundamental teaching practices for

physical education students. From this study, it was found that teaching skills increased progressively during students participated in the internship ($F(1.416) = 6,658, p < 0.01, \eta^2 = 0.117$), unfortunately, the significant increase was only found in measurement 1 to measurement 2 ($\Delta = 0.021, p < 0.05$), while measurement 2 to measurement 3 was not found to be significant ($\Delta = 0.008, p > 0.05$). The learning effectiveness itself does not develop during the internships ($F = 1.173, p > 0.05, \eta^2 = 0.316$). The technical factors for data collection and the short duration of the apprenticeship were considered the possible causes. For this reason, the MBKM-based internship raises new hopes in providing scientific development and a real experience to students by promising a longer duration of internship implementation. However, as a consequence, higher education institutions must prepare good management in setting goals through an integrated assessment instrument of all subjects included in the apprenticeship program. Besides, the provision of human resources in the form of lecturers for each subject should be apprenticed. Also, there are many highly qualified tutors, schools that have superior qualifications, and an incentive system or other forms that are appropriate and commensurate with the performance of civil service teachers in assisting students during their internships.

REFERENCE

- Barney, D., Christenson, R., & Pleban, F. (2012). Pre-service physical education teachers' knowledge of appropriate instructional practices in secondary school physical education. *Journal of Teacher Education for Sustainability, 14*(1), 30–38. <https://doi.org/10.2478/v10099-012-0003-y>
- Barney, D., & Pleban, F. (2006). Pre-service physical education teacher's perceptions of teaching before and after a semester long elementary physical education practicum experience. *Physical Educator, 63*(1), 46–52. <https://js.sagamorepub.com/pe/article/view/2406>
- Bergsmann, E., Schultes, M.-T., Winter, P., Schober, B., & Spiel, C. (2015). Evaluation of competence-based teaching in higher education: From theory to practice. *Evaluation and Program Planning, 52*, 1–9. <https://doi.org/10.1016/j.evalprogplan.201>

- 5.03.001
Bert, G., & Richards, K. A. R. (2018). Developing a future teachers' club to facilitate physical education teacher recruitment: Do you hear the call? *Journal of Physical Education, Recreation & Dance*, 89(7), 9–15. <https://doi.org/10.1080/07303084.2018.1490224>
- Clifford, M., Behrstock-Sherratt, E., & Fetters, J. (2012). The ripple effect: A synthesis of research on principal influence to inform performance evaluation design. In *American Institutes for Research*. American Institutes for Research. <http://www.nysed.gov/common/nysed/files/principal-project-the-ripple-effect.pdf>
- Dhuey, E., & Smith, J. (2014). How important are school principals in the production of student achievement? *Canadian Journal of Economics/Revue Canadienne d'économique*, 47(2), 634–663. <https://doi.org/10.1111/caje.12086>
- Faikhamta, C., & Clarke, A. (2018). Thai cooperating teachers' motivations and challenges in supervising student teachers during their internship program. *Kasetsart Journal of Social Sciences*, 1–7. <https://doi.org/10.1016/j.kjss.2017.12.018>
- Fransson, G., & K. McMahan, S. (2013). Exploring research on mentoring policies in education. *International Journal of Mentoring and Coaching in Education*, 2(3), 218–232. <https://doi.org/10.1108/IJMCE-05-2013-0031>
- Ganefri, G. (2016). Kolaborasi strategi pemberdayaan lintas institusi dan participatory management menuju sistem rekrutmen dan distribusi guru yang proporsional-efektif di Indonesia. *Konvensi Nasional Pendidikan Indonesia (KONASPI) VIII Tahun 2016*, 35–40.
- Gelman, A., & Hennig, C. (2017). Beyond subjective and objective in statistics. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(4), 967–1033. <https://doi.org/10.1111/rssa.12276>
- Guo, Y., Logan, H. L., Glueck, D. H., & Muller, K. E. (2013). Selecting a sample size for studies with repeated measures. *BMC Medical Research Methodology*, 13(1), 100. <https://doi.org/10.1186/1471-2288-13-100>
- Hasegawa, E., Takahashi, T., Urai, T., & Matsumoto, T. (1995). Development of an instrument for formative evaluation and its diagnostic standard of physical education class in elementary school. *Japanese Journal of Sport Education Studies*, 14(2), 91–101. <https://doi.org/10.7219/jjses.14.91>
- Hill, G., & Brodin, K. L. (2004). Physical education teachers' perceptions of the adequacy of university coursework in preparation for teaching. *Physical Educator*, 61(2), 75–87. <https://js.sagamorepub.com/pe/article/view/2584>
- Imberman, S. A., & Lovenheim, M. F. (2015). Incentive strength and teacher productivity: Evidence from a group-based teacher incentive pay system. *Review of Economics and Statistics*, 97(2), 364–386. https://doi.org/10.1162/REST_a_00486
- Khory, F. D., Hartoto, S., Sudarso, S., & Prakoso, B. B. (2018). Persepsi guru pamong terhadap kualitas mahasiswa calon guru PJOK saat melaksanakan program pengelolaan pembelajaran di sekolah mitra. *Jurnal SPORTIF : Jurnal Penelitian Pembelajaran*, 4(1), 38. https://doi.org/10.29407/js_unpgri.v4i1.12092
- Marcon, D., Graça, A. B. S., & Nascimento, J. V. (2012). Analysis of the pedagogical knowledge of prospective physical education teachers. *Kinesiology*, 44(2), 113–122. <https://hrcak.srce.hr/file/139182>
- Marli, S. (2012). Supervisi klinis bagi calon guru dalam program pengalaman lapangan (PPL). *Jurnal Visi Ilmu Pendidikan*, 5(2), 432. <https://doi.org/10.26418/jvip.v5i2.58>
- Maulipaksi, D. (2016). *7 provinsi raih nilai terbaik uji kompetensi guru 2015*. Laman Kementerian Pendidikan Dan Kebudayaan Indonesia.
- Meyers, L. S., Gamst, G. C., & Guarino, A. J. (2013). *Performing data analysis using IBM SPSS*. John Wiley & Sons, Inc.
- Muhtar, T., & Dallyono, R. (2020). Character education from the perspectives of elementary school physical education teachers. *Cakrawala Pendidikan*, 39(2), 395–408. <https://doi.org/10.21831/cp.v39i2.30647>

- Nagro, S. A., DeBettencourt, L. U., Rosenberg, M. S., Carran, D. T., & Weiss, M. P. (2017). The effects of guided video analysis on teacher candidates' reflective ability and instructional skills. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, 40(1), 7–25.
<https://doi.org/10.1177/0888406416680469>
- Rikard, G. L., & Knight, S. M. (1997). Obstacles to professional development: Interns' desire to fit in, get along, and be real teachers. *Journal of Teaching in Physical Education*, 16(4), 440–453.
<https://doi.org/10.1123/jtpe.16.4.440>
- Şen, A. I. (2010). Effects of peer teaching and microteaching on teaching skills of pre-service physics teachers. *Egitim ve Bilim*, 35(155), 78–88.
<https://app.trdizin.gov.tr/makale/T1RRNU1qSXk/effects-of-peer-teaching-and-microteaching-on-teaching-skills-of-pre-service-physics-teachers>
- Shen, J., Leslie, J. M., Spybrook, J. K., & Ma, X. (2012). Are principal background and school processes related to teacher job satisfaction? A multilevel study using schools and staffing survey 2003-04. *American Educational Research Journal*, 49(2), 200–230.
<https://doi.org/10.3102/0002831211419949>
- Suroto, S. (2005). *Examining the relationship among studnets' phisical activity level, students' learning behaviors, and students' formative class evaluation during elementary school physical education classes*. University of Tsukuba.
- Suroto, S., & Khory, F. D. (2015). *Peningkatan keterampilan mengelola pembelajaran siswa aktif melalui pendekatan lesson study (Studi pada guru PJOK SDN di Kecamatan Taman Sidoarjo)*.
- Suroto, S., & Takahashi, T. (2005). Students' physical activity level, students' learning behavior, and their formative class evaluation during fitness units of elementary school physical education classes. *International Journal of Sport and Health Science*, 3, 10–20.
<https://doi.org/10.5432/ijshs.3.10>
- Tohir, M. (2020). *Merdeka Belajar: Kampus Merdeka*. OSF Preprints.
<https://doi.org/10.31219/osf.io/sv8wq>
- Ward, P. (2013). The Role of content knowledge in conceptions of teaching effectiveness in physical education. *Research Quarterly for Exercise and Sport*, 84(4), 431–440.
<https://doi.org/10.1080/02701367.2013.844045>
- Ward, P. (2016). Policies, agendas, and practices influencing doctoral education in physical education teacher education. *Quest*, 68(4), 420–438.
<https://doi.org/10.1080/00336297.2016.1234964>
- Wijaya, A. M., & Astono, A. (2006). *Uji coba instrumen baku evaluasi pembelajaran dikdasar di SLTP negeri se Kota Sudaranaya*.