

Evaluating the learning goal attainment in Nihongo Noryoku Shiken N5 course

Ai Sumirah Setiawati*; Wagiran; Subyantoro

Universitas Negeri Semarang, Indonesia

*Corresponding Author. E-mail: ai.sumirah@mail.unnes.ac.id

ARTICLE INFO

Article History

Submitted:

13 August 2021

Revised:

13 October 2021

Accepted:

6 December 2021

Keywords

Evaluation; learning goal;
Nihongo Noryoku Shiken;
N5

Scan Me:



ABSTRACT

The objective of the study is to evaluate the learning implementation of the *Nihongo Noryoku Shiken N5* course as a new course offered in the 2020 curriculum at the Japanese Language Education Study Program in UNNES. A goal-oriented evaluation model by Tyler was employed in the evaluation process to measure how far the learning goals that have been determined at the beginning of the lecture are achieved. This evaluative research relied on a quantitative descriptive approach; its steps comprised (1) data collection, (2) data processing and analysis, and (3) recommendation formulation. According to the evaluation results, the learning goals of *Nihongo Noryoku Shiken N5* had been met, although it was not that significant. Several conditions cause such issues. (1) In the class, some students had good Japanese language skills; they claimed that the lesson was too easy, so their progress was not significant. (2) The evaluation instruments were not valid. (3) Other conditions, e.g., the implementation of online learning and tests, hindered classroom activities and monitoring. Improving the teaching and learning materials and evaluation instruments is recommended. Furthermore, some additions to the curriculum need to be considered, such as adding the credits for the grammar or *Bunpo* course.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



How to cite:

Setiawati, A., Wagiran, W., & Subyantoro, S. (2021). Evaluating the learning goal attainment in Nihongo Noryoku Shiken N5 course. *Jurnal Penelitian dan Evaluasi Pendidikan*, 25(2), 151-161.

doi:<https://doi.org/10.21831/jpep.v25i2.43128>

INTRODUCTION

Evaluation is among the stages that are central to learning activities. The reason lies in its function to provide target or learning goal attainment information. The evaluation also describes the lack or the advantage of learning activities. Improvement, such as revising some points in the lesson plan, will be the next step if some issues are identified in the evaluation result. In other words, evaluation can serve as a reference for proposing recommendations to stakeholders or educators based on the data of the achievement or progress, obstacles, and other essential information. Such information will be beneficial in determining the next step in advancing the learning activities.

Evaluation programs can be defined as a systematical operation formed by many complex processes involving data collection, observation, and analysis, culminating in assessment outputs (components in the outputs have been scrutinized) of evaluated learning (Mizikaci, 2006, p. 41). Besides, evaluation is defined as a means to obtain a valid and reliable assessment based on the measurement types (qualitative or quantitative) that provide actual results.

As a policymaker of a program or learning, educators are responsible for evaluating activities managed by themselves. In that process, they can also refer to multiple evaluation publications in the education field. All evaluation processes by other parties can serve as one of the resources in designing, implementing, and evaluating a program or learning activity. The reference can be applied as a whole or modified based on the institution or students' charac-

teristics. Nevertheless, evaluating the reference is needed to examine the effectiveness of the reference. This process also enables one to determine whether or not the inputs from others are suitable for their class and identify approaches required for attaining the learning goals.

One example of an evaluation instrument for problem-based learning is seen in a study by [Hinoto et al. \(2017\)](#) that proposes a quantitative evaluation based on the capability maturity integration model. Another evaluation model called CIPP (context, input, process, and product) has been proposed by [Pramesti \(2020\)](#); this model was implemented in the mathematics subject at boarding schools. Further, CIPP is applicable in web-based collaborative learning with multiple intelligence ([Tiantong & Tongchin, 2013](#)). Besides, evaluation is implemented in a workshop or practice-based subjects, including bilingual mathematics and natural science training for teachers ([Raihanati et al., 2015](#)) and music training in Pontianak ([Gery et al., 2018](#)).

In language education, research on evaluation in Indonesia has been extensively conducted in the teaching and learning process of Arabic, English, and Indonesia. The publications of evaluation results in the form of an article are also displayed, as follows: First, in the field of the Arabic language, [Mashur and Baili \(2020\)](#) evaluated the Arabic language program in Al-Azhar Islamic boarding school and Islamic Studies College in Kudus on seven essential elements in education: input (students), materials (curriculum), lecturers, method (educative interaction), infrastructure, human environment, and non-human environment. Second, in the field of the English language, [Bendriyanti and Dewi \(2014\)](#) performed an in-depth evaluation towards the implementation of the English language course curriculum at Dehasen University, Bengkulu. This evaluation was carried out using the countenance evaluation model developed by Stake. Another sample is a study by [Ting et al. \(2017\)](#) that explored a telepresence transformation and its relation with the learning process. Third, in the field of the Indonesian language, [Aryanika \(2015\)](#) evaluated the Indonesian language learning programs in universities so that the students can develop their skills in using the Indonesian language correctly and adequately, both in oral skills and writing skills.

In particular, research on evaluation in Japanese language learning has also been conducted ([Cuī, 2013](#); [Kano & Wei, 2014](#); [Lee, 2018](#); [Qiao, 2015](#); [Suematsu, 2017](#); [Suzuki & Fujimori, 2014](#); [Yokota, 2014](#)). As the authors assert, evaluating the education program and learning process that applies a specific Japanese language model is rare. In general, evaluations have been carried out, as in the following samples: evaluation of the use of specific languages and behaviors in the *Kaima* (the Japanese language speaking) program ([Lee, 2018](#)). In this context, the evaluation refers to the speaker and the interlocutor in a conversation. For instance, the use of the adjective “*utsukushii*” (“beautiful” in English) is the evaluation of the utterance spoken by the interlocutor. Besides diction, evaluation is also represented by intonation, mimic, and other parts. Thereby, the difference between such an evaluation and program or learning process evaluation is obvious.

Moreover, [Cuī \(2013\)](#) and [Qiao \(2015\)](#) evaluated both native and non-native teachers towards Japanese language learners. In the study, Cui elaborated on the causal relationship between evaluation conducted by native and non-native Japanese language teachers. It was reported that native and non-native Japanese language teachers have a difference of perspective in performing the evaluation. Unlike Cui, Qiao only researched the evaluation done by native teachers without comparing it to the non-native teachers. The evaluation was carried out on several aspects of speaking, such as discourse skills, sociolinguistic skills, negotiation strategy skills, grammar, and vocabulary expression. Even though the studies are different, they conduct research on the speaking skill learning program evaluations.

Further, studies on self-evaluation ([Suzuki & Fujimori, 2014](#); [Yokota, 2014](#)) have also been conducted in Japanese language learning. Meanwhile, another common conduct in the evaluation scholarship is survey-based evaluation ([Suematsu, 2017](#)), or test-based evaluation ([Kano & Wei, 2014](#)) that uses TTBJ (Tsukuba Test Battery of Japanese), a Japanese language

test researched and developed for years in the Student Center of Tsukuba University. Such a test system allows the student to partake in a web-based test; the test was already released in public for individual examination in June 2013.

As have been stated priorly, studies on Japanese learning evaluation apply specific evaluation models, namely experimental and quasi-experimental model, Kirkpatrick four-level model, logic model, and CIPP model (Frye & Hemmer, 2012), or evaluation model as stated by Muryadi (2017), i.e., discrepancy model, CIPP model, responsive evaluation model, formative-summative evaluation model, measurement model, and goal-free evaluation model. Evaluation of programs or learning process that applies a particular model, particularly the goal-oriented evaluation model (Mardiah & Syarifudin, 2019), as far as the author's search, has not been found in research on evaluation of Japanese language learning. The goal-oriented evaluation was developed by Ralph W. Tyler, who was born on April 22, 1902, in Chicago. Fitzpatrick (Nurman, 2016, p. 209) asserts that Tyler established seven steps to determine the extent of goal achievement of a program/learning process: (1) determining general objectives, (2) classifying the objectives, (3) defining the objectives in behavior contexts, (4) determining the situation in which the goal attainment can be demonstrated, (5) developing or selecting measurement techniques, (6) collecting performance data, and (7) comparing performance data and behavioral data that depict the objectives.

Based on the rationale, the author intends to apply the model in Japanese learning skills, especially in the learning process aimed at training Japanese language skills in preparation for taking the JLPT (Japanese Language Proficiency Test). The study focuses on the *Nibongo Noryoku Shiken N5* course given to the semester II students. The intention is that the students can master the Japanese language skills at N5 level, specifically in topics of letters and vocabulary, grammar and reading, and listening. This is a new program incorporated into the new curriculum in the Japanese Education Department, Universitas Negeri Semarang (UNNES). This curriculum combines Japanese receptive language mastery with productive Japanese language mastery to complement each other.

The *Nibongo Noryoku Shiken N5* course is presented in the new curriculum to address the problem of students' unpreparedness in facing the Japanese language proficiency test because, since 2018, The Department of Japanese Education in UNNES has implemented a policy to use the *Marugoto* coursebook which prioritizes the productive Japanese language mastery. This book is different from the previous book referred, i.e., *Minna no Nibongo* (Honsatsu edition) that many regarded as capable of training the receptive Japanese language mastery that allows the student to prepare themselves before partaking in the JLPT.

As a new program and curriculum, we need to know how far this effort can help students improve their Japanese ability, as a goal of this program, especially in partaking in the JLPT. If the effort is successful, it is hoped that it can provide a reference to other universities in Indonesia that have been hesitant to apply the Japan Foundation Standard (Fujinaga & Yūki, 2013; Hiromi et al., 2012) using the *Marugoto* coursebook because there are concerns that students' graduation rate in following the JLPT will be low. Besides, the JLPT issue is still very important since ownership of a JLPT certificate is still the main requirement, for example, to become a Japanese language teacher, work in a Japanese company in Indonesia, or for those who will take part in an internship program in Japan as a caregiver or *tokutei ginou* (trainee with specific skill), which in recent years Japan has been doing massive recruitment.

Therefore, it is hypothesized that the preparation course for the JLPT can assist the students in achieving the N5 level in Japanese language mastery. In this regard, the evaluation result is expected to be a reference in the Japanese language teaching field. Practically, this study is also expected to be a reference for the Department of Japanese Education stakeholders to evaluate the Japanese language education program; to achieve the optimum quality and increase the accreditation of the Department of Japanese Education of UNNES.

RESEARCH METHOD

This evaluative research relied on a quantitative descriptive approach. Evaluative research mainly focuses on the final recommendation, which states that an object of evaluation can be maintained, improved, or even discontinued in line with the data obtained (Rozak, 2018, p. 173). On that ground, the evaluation process in this study comprised three steps, namely (1) data collection, (2) data processing and analysis, and also (3) recommendation formulation.

The primary data were collected using the pre-and-post-test technique, as exemplified by Tyler, who developed the goal-oriented evaluation model as a measurement technique. Such a technique aims to observe the changes that occurred in individuals, activities, or programs and measure the extent of the changes (Nurman, 2016, p. 209). We use validity content on this instrument by utilizing the JLPT questions that have been compiled and published by The Japan Foundation. Then, the reliability was confirmed by using IBM SPSS Statistics 24 version. In addition, documentation was also employed to identify any required secondary data or information to support the study, such as the document of the semester lesson plan, learning materials, and other evaluation-related data.

FINDINGS AND DISCUSSION

Goal-oriented evaluation is employed to observe the achievement of goals. The description of the characteristic of the course, students, learning objectives, and the change in students' ability is elaborated as follows.

Course Characteristics

The *Nibongo Noryoku Shiken* N5 course (the name of the course pronounced with a long O sound is not stated in the *nooryoku* word spelling) is a new course program in the Department of Japanese Education UNNES. Incorporating this course in the newly applied curriculum in the students of the class of 2020 refers to the consideration of learning and discussion of JLPT questions. This course is to be implemented in the second semester; the higher the semester, the higher the mastery level. For instance, if at the second semester the level is N5, the third semester will begin at N4 level, and so forth, until the mastery level reaches N1 as the highest mastery level.

Five learning objectives have been proposed based on the semester's lesson plan, including: (1) students can master *hiragana*, *katakana*, and N5 *kanji*, (2) students can master N5 Japanese vocabularies, (3) students can master N5 Japanese grammar, (4) students demonstrate listening skills at the N5 level, and (5) students demonstrate reading comprehension skills at the N5 level. The course consists of 16 meetings that comprise 14 meetings discussing material, one meeting for mid-term examination, and one meeting for final-term examination (Universitas Negeri Semarang, 2008). The materials are the old JLPT questions that are divided into several phases by referring to the types, as elaborated in Table 1.

Table 1. Lessons and Number of Meetings

No.	Lessons	Number of Meetings
1.	Letter	1
2.	Vocabularies	1
3.	Grammar	4
4.	Reading	4
5.	Listening	4
	Total	14

The learning process begins with a pre-test, then, at each meeting, a quiz is given by referring to the material. The course ends with a post-test which can also be utilized as scores for final examination that must be inputted into the academic system at UNNES.

The pre-test and quiz are discussed right after the students finish both of the tests. Moreover, the students are asked to try by any means they can to solve the problems prior to providing guidance at each question. The quiz form is chosen to familiarize the students with the real JLPT questions.

Student Characteristics

A total of 75 students enrolled in the *Nihongo Noryoku Shiken* course; they were divided into two classes. Some students missed several meetings in progress, and even one student did not attend the class from the first meeting. In the end, 72 students had an attendance percentage of more than 75%, the minimum requirement to pass the course.

As many as 77% of students have learned Japanese before entering the university. They learned Japanese at school or enrolled in a course; some even learned the language themselves. Additionally, the survey revealed that 13% of 75 students had taken the JLPT, mostly N5 level. Moreover, one student had taken the N3 level.

Based on the description of the student characteristics, it can be seen that most of the students have initial Japanese language skills at the N5 level. However, in the preliminary survey, although the pretest was carried out, there was no detailed review of their abilities. To find out the detailed abilities of the teaching students, the lecturer can conduct a questionnaire study on "JLPT Can-do" using JLPT Can-do Statements (Naganuma, 2008, p. 51) as has been done by Yanagisawa (2018) and Suzuki and Fujimori (2014). Then, the JLPT training materials or modules are adjusted to the results of the questionnaire (Naganuma, 2009). Thus, the material provided during the course is in accordance with the needs or goals of increasing students' Japanese language skills.

Learning Process

Nihongo Noryoku Shiken N5 is an online course taught using Elena, a moodle-based LMS (learning management system) by UNNES. The platform allows students to study both synchronously and asynchronously. The students can discuss, read materials, submit assignments, and work on quizzes using the platform. Synchronous meetings are also conducted via Zoom. Other learning media include Google Forms (for quizzes and tests) and WhatsApp group (for class discussions and instructions).

In terms of active learning using LMS or other media such as google quizzes, it can be said that students are very active because only 7% of the average inactivity is recorded during one semester of lectures. This shows that students can carry out the values of independent character and responsibility according to what is mandated in achieving learning outcomes.

However, in synchronous lectures using the zoom meet facility, almost all students are reluctant to open their cameras, and only a small number are actively asking or volunteering to answer or discuss the material being studied. Most of the others are just waiting to be called by name to answer questions both posed by the instructor and questions on quiz questions.

Then, judging from the average daily quiz scores (see Table 2), the lowest score is 61.72, and this value is included in the sufficient category. On the other hand, the highest average that can be achieved by students is 89.53, and this value is included in the very good category based on the academic guidelines that apply at UNNES. Thus, it can be said that no one has failed in learning Japanese in this *Nihongo Noryoku Shiken* course because the minimum score that is regulated in the academic policy at UNNES is grade C or sufficient (Universitas Negeri Semarang, 2008).

Table 2. Daily Quiz Score Recap

Aspect	Av.	HS	LS	Int.
Letter	89.53	56	92	36
Vocabularies	61.72	27.5	95	67.5
Grammar	65.81	17.65	89.34	71.69
Listening	89.16	12.5	100	87.5
Reading	82.5	41.16	97.5	56.34
All Aspects	77.65	61.72	89.53	27.81

Annotation: Av. : Average score of the class
 HS : Highest score
 LS : Lowest score
 Int. : Interval

Moreover, considering the overall average daily quiz score, which is 77.65 and is included in the Good category, it can be said that the overall learning process has been going well. Another thing that can be concluded from Table 2, results of student daily quizzes reported that the average student scores were in a good category. Still, this research identifies gaps among the score of each learning goal (mastery of *N5 hiragana*, *katakana* and *kanji*, N5 vocabularies, N5 grammar, N5 listening, and N5 reading comprehension).

From Table 2, the vocabulary aspect scores the lowest (61.72), and the highest is the mastery of Japanese letters (89.53). The table also reveals the lowest and highest score interval, meaning that the data are not normally distributed. A measurement result is classified good if the data are distributed normally (Akbar & Anhar, 2018, p. 72). On that ground, this research proceeded to the normality test. This test determines whether the collected data are normally distributed or retrieved from the average population (Fahmeyzan et al., 2018, p. 32).

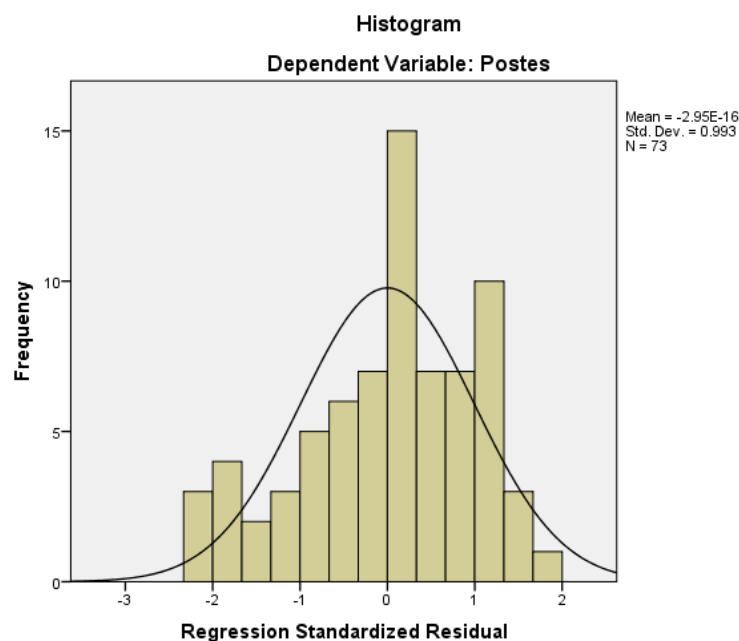


Figure 1. Normality Test Graph

Based on Figure 1, the data are not distributed normally. This confirms the necessity to re-examine the evaluation instrument or the tests. In the final analysis regarding the evaluation of the learning process, the learning activities were well-implemented. Notable issues revolve around connection problems in some students during working on quizzes or during the regular class activity using Zoom; some lecturers also had the same connection problem in giving lectures.

Student Progress

Nibongo Noryoku Shiken N5 intends to help students master Japanese language skills at the N5 level, covering the aspects of letters, vocabulary, grammar, listening, and reading. The course is designed to prepare students for taking JLPT (Japanese Language Proficiency Test).

Based on the pre-and post-test results, the students' skills were improved. The average score of the pretest is 77.32, and the post-test is 84.77. It can be concluded that there has been an increase in students' Japanese language skills. Then, in the middle of the semester, a test was also carried out, the results showed that the average student score was 83.39. The addition of the average score from 77.32 to 83.39 and then 84.77 shows that students' Japanese language skills continue to improve, not fluctuate. Although the increase is not very high, especially from midtest to post-test scores, this is good to increase or change abilities. However, the progress is yet considered significant, as seen in the t-test result in Table 3. This is due to a decline in the individual score of each student.

Table 3. T-Test Result

Coefficients ^a						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	66.456	6.017		11.044	.000
	Pre-test	.231	.076	.341	3.055	.003

^aDependent Variable: Post-test

From the test results, the present research analyzed secondary data, i.e., learning materials, semester lesson plans, and survey results. First, the objective of learning stated in the lesson plans was examined. It is revealed that the learning goals focus on Japanese language skills and learners' attitudes or characters. Since the test was administered online, it was difficult to ensure whether or not the students demonstrated honesty. Thus, the validity of the students' skills is not comprehensive. The lecturers had confirmed cheating practices. The inability to attain valid results of students' overall progress also blames the materials that are considered too easy by some students.

Second, the survey shows that 77% of students had learned Japanese language. In fact, this information reveals that most students' Japanese language mastery is considered good (Universitas Negeri Semarang, 2008). The average pre-test score reached 77.32.

Despite the number of students with high Japanese language skills, and they also show improvement in the final or post-test, it cannot be said that this study's objectives or learning goals have been fully achieved. The prepost-test measurement method is one way that can be done in evaluating learning with a goal-oriented model. In general, this evaluation model also has a weakness: the lack of a real evaluation component, emphasizing measuring the achievement goals rather than the worth of the goals themselves (Putra, 2012, p. 67).

As explained at the end of the "student characteristic" section, at the beginning of learning, it is necessary to carry out a detailed and concrete evaluation of the initial abilities of students. At the end of the lesson, the same thing is done. Students re-evaluate themselves with the self-check format on the JLPT Can-do List or Can-do Statements (Yanagisawa, 2019). Thus, we will get more complete information about the achievement of learning goals of the *Nibongo Noryoku Shiken* course in more comprehensive and not just statistical information.

Besides the test, the lecturer can conduct interviews to find out the students' ability level of the Japanese Language. As Ogasawara (2010) did in his research to find out whether the classification judgment by the interview test of the international student special course that has been conducted so far can appropriately judge the Japanese language ability that the learner

has acquired. The result said that the interview technique is effective for judging the Japanese language ability of the student.

Recommendation for Japanese Language Education Study Program, UNNES

As stated earlier in the introduction section, the Japanese Language Education study program has decided to use *Marugoto* (Kawashima, 2015) as the textbook for the *Nihongo Noryoku Shiken N5* course since 2018. The book emphasizes practices on communicative Japanese language skills (The Japan Foundation, 2017). It should be noted that the difficulty of the materials of the book is not designed systematically (i.e., from the easiest to hardest). Grammar contents are incorporated in the book; the lessons, however, are delivered using a communicative approach. The difficulty of each grammar topic is not listed systematically. For example, complex grammar, such as the verb form *-te + imasu* is taught earlier in the book (in chapter 4). On the contrary, in another similar book, i.e., *Minna no Nihongo*, the same topic will be discussed later (in chapter 15).

Minna no Nihongo applied an audio-lingual approach with a focus on grammar mastery. In the book, the difficulty of the grammar topic is delivered from the simple to the complex level. The new curriculum of the Japanese Language Education study program applies the combination of two approaches of the two books. *Marugoto* became a primary textbook for 4-unit courses, while *Minna no Nihongo* is mainly for two-unit courses.

Provided in the following is the recommendation for the department regarding the use of the books in learning activities. (1) *Marugoto* should be used in the two-credits course. The focus of *Marugoto* on communication activities enables lecturers to incorporate more topics for the speaking course with *kaiwa* (speaking) course. Another issue is courses with limited contact hours, such as *bunpo* or grammar (two credits). (2) *Minna no Nihongo* can be used as the textbook for courses with four credits. *Marugoto* is used in lectures with four credits in the current curriculum, and vice versa *Minna no Nihongo* is used in lectures with two credits. This is proposed with the consideration that the *bunpo* course, which currently only has two credits, is an important course that provides knowledge input to students to be applied to Japanese speaking or communication courses. (3) To further advance the learning activities and enhance students' language skills, other books can supplement *Minna no Nihongo*. It should be noted that the supplement must be verified first by the department. (4) Decision-makers or the department should take into account whether or not to separate students with zero Japanese language skills and students with a learning experience in the Japanese language. This is based on the condition that some freshmen have already taken JLPT, ranging from N5, N3, to N2. (5) Lecturers of *Nihongo Noryoku Shiken* should better the evaluation instruments to ensure the validity of students' progress. Lecturers of *Nihongo Noryoku Shiken* should improve some lessons or topics discussed in the course.

CONCLUSION

According to the evaluation results, the learning goals of *Nihongo Noryoku Shiken N5* had been met. This is evident from students' learning progress, although it is not that significant due to some factors: (1) some students already have good Japanese language skills, (2) the lessons are deemed too easy, thus resulting in less significant learning progress, (3) invalid evaluation instruments, (4) ineffective online learning, which hinders the evaluation and monitoring of students, (5) ineffective online learning, which hinders the evaluation and monitoring of students, and (6) ineffective online learning, which hinders the evaluation and monitoring of students. Recommendations have been proposed to improve the quality of the Japanese Language Education Study Program at UNNES. It is expected that the recommendations will help stakeholders formulate policies for better learning activities and graduates' quality.

REFERENCES

- Akbar, S., & Anhar, W. (2018). Kajian hasil pengukuran Undercarriage Bulldozer Komatsu D375A-5 di PT. Pama Persada Nusantara site Batukajang. *JST (Jurnal Sains Terapan)*, 4(1), 70–75. <https://doi.org/10.32487/jst.v4i1.455>
- Aryanika, S. (2015). Evaluasi pembelajaran Bahasa Indonesia di perguruan tinggi. *Al-Idarah: Jurnal Kependidikan Islam*, 5(1), 116–130. <http://ejournal.radenintan.ac.id/index.php/idaroh/article/view/757>
- Bendriyanti, R. P., & Dewi, C. (2014). Model “Countenance Stake” dalam evaluasi pembelajaran Bahasa Inggris di perguruan tinggi. *Prosiding Seminar Nasional Riset Inovatif II*, 134–139. http://digilib.mercubuana.ac.id/manager/tl@file_artikel_abstrak/Isi_Artikel_587541717147.pdf
- Cuī, W. (2013). Nihongo gakushūsha no hatsuwa ni taisuru Nihongo bogo washa no hyōka: Nihongo kyōshi to hi Nihongo kyōshi no inga moderu o chūshin ni. *Kokuritsu Kokugo Kenkyūsho Ronshū*, 5, 1–26.
- Fahmeyzan, D., Soraya, S., & Etmy, D. (2018). Uji normalitas data omzet bulanan pelaku ekonomi mikro Desa Senggigi dengan menggunakan Skewness dan Kurtosi. *Jurnal V/ARLAN*, 2(1), 31–36. <https://doi.org/10.30812/varian.v2i1.331>
- Frye, A. W., & Hemmer, P. A. (2012). Program evaluation models and related theories: AMEE Guide No. 67. *Medical Teacher*, 34(5), e288–e299. <https://doi.org/10.3109/0142159X.2012.668637>
- Fujinaga, K., & Yūki, N. (2013). JF Nihongo kyōiku sutandādo o riyō shita `kyōshi-muke nihongo kōza` kaizen no kokoromi. *Kokusaikōryūkin Nihongo Kyōiku Kijō*, 9, 89–107.
- Gery, G., Mering, A., & Silaban, C. Y. (2018). Evaluasi program pelatihan musik di Pontianak. *Jurnal Pendidikan Dan Pembelajaran Khatulistiwa (JPPK)*, 7(10), 1–10. <https://jurnal.untan.ac.id/index.php/jpdpb/article/view/29013>
- Hinoto, N., Itō, M., & Ōba, M. (2017). Nōryoku seijukudo moderu tōgō ni motodzuita PBL ni okeru teiryōteki gakushū hyōka shuhō no teian. *Nihon Sofutowea Kagakukai*, 34, 1–7.
- Hiromi, K., Tomoyo, S., & Naomi, H. (2012). JF Nihongo kyōiku sutandādo junkyo kōsubukku no kaihatsu. *Kokusaikōryūkin Nihongo Kyōiku Kijō*, 8, 103–117.
- Kano, C., & Wei, W. (2014). Gaikokujin Nihongo gakushūsha no kanji-ryoku no hyōka ni tsuite: TTB (Tsukuba Nihongo Tesuto-shū) o riyō shite. *JSL Kanji Gakushū Kenkyūkai-Shi*, 6, 54–62.
- Kawashima, K. (2015). Nihongo gakushū saito “Marugoto + (Marugoto Purasu)” no kaihatsu—Kadai suikou to ibunka rikai o tasukeru Website. *Kokusai Kouryū Kikin Nihongo Kyōiku Kijō*, 11, 37–52. <http://www.jpf.go.jp/j/project/japanese/teach/research/report/11/pdf/03.pdf>
- Lee, S. (2018). Ryūgakusei to nihonjin chūta no gakushū katsudō: Kaiwa ni okeru kyōka ni chakumoku shite. *Gengo Bunka Kyōiku Kenkyū*, 16, 157–176.
- Mardiah, M., & Syarifudin, S. (2019). Model-model evaluasi pendidikan. *MITRA ASH-SHIBYAN: Jurnal Pendidikan Dan Konseling*, 2(1), 38–50. <https://doi.org/10.46963/mash.v2i1.24>

- Mashur, I. Ihatif, & Baili. (2020). Evaluasi pelaksanaan program bahasa Arab di Pondok Pesantren Al-azhar Malang. *Tadbir: Jurnal Manajemen Pendidikan Islam*, 8(1), 39–52. <https://doi.org/10.30603/tjmpi.v8i1.753>
- Mizikaci, F. (2006). A systems approach to program evaluation model for quality in higher education. *Quality Assurance in Education*, 14(1), 37–53. <https://doi.org/10.1108/09684880610643601>
- Muryadi, A. D. (2017). Model evaluasi program dalam penelitian evaluasi. *JURNAL ILMIAH PENJAS (Penelitian, Pendidikan Dan Pengajaran)*, 3(1), 1–16. <http://ejournal.utp.ac.id/index.php/JIP/article/view/538>
- Naganuma, N. (2008). *Can - Do shakudo wa ikani eigo kyōiku o henkaku shi uru ka: Can - Do kenkyū no bōkō-sei*.
- Naganuma, N. (2009). Can - do hyōka — gakushū tasuku ni motodzuku mojūru-gata shirabasu kōchiku no kokoromi. *Tōkyōgaikokugodaigaku Ronshū*, 79, 87–105.
- Nurman, M. (2016). Evaluasi program pendidikan: "Pendekatan evaluasi program berorientasi tujuan (goal-oriented evaluation approach: Ralph W. Tyler)". *El-Tsaqafah: Jurnal Jurusan PBA*, 15(2), 203–212. <https://journal.uinmataram.ac.id/index.php/eltsaqafah/article/view/254>
- Ogasawara, N. (2010). Intabyū ni yoru Nihongo nōryoku sutēji hyōka-hō no kenshō — Dankai betsu shitsugi ōtō ni yoru mensetsu shiken hyōka to hikki shiken hyōka no kanren. *Shakai Jōbō Ronso*, 14, 31–49.
- Pramesti, S. L. D. (2020). Evaluasi pembelajaran Matematika pada Boarding School berdasarkan model CIPP. *Integral: Pendidikan Matematika*, 11(1), 17–32. <https://ejournal.umc.ac.id/index.php/JNR/article/view/1139>
- Putra, A. T. A. (2012). Evaluasi program pendidikan: "Pendekatan evaluasi program berorientasi tujuan (goal-oriented evaluation approach: Ralph W. Tyler)." *Shautut Tarbiyah*, 18(1), 55–68. <https://ejournal.iainkendari.ac.id/shautut-tarbiyah/article/view/65>
- Qiao, X. (2015). Taiwanjin Nihongo gakushūsha no bijinesu kaiwa ni mirareru tokuchō — Nihongo bogo washa no hyōka kara. *Hitotsubashi Daigaku Kokusai Kyōiku Sentā Kiyō*, 6, 65–78.
- Raihanati, R., Supriyati, Y., & Rahayu, W. (2015). Efektivitas program pendidikan guru MIPA kelas bilingual Universitas Negeri Jakarta. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 01(2), 27–32. <https://doi.org/10.21009/1.01205>
- Rozak, A. (2018). Modernisme pembelajaran Bahasa Arab berbasis pesantren di Rangkasbitung Banten. *Arabi: Journal of Arabic Studies*, 3(2), 167–180. <https://doi.org/10.24865/ajas.v3i2.110>
- Suematsu, D. (2017). "Atarashī Nihongo Gakushūsha" no jittai to gakushū komyuniti ni taisuru hyōka Feisubukku gurūpu "The Nihongo Ran'ningu Komyuniti" de no chōsa kekka kara. *Gengo Bunka Kyōiku Kenkyū*, 15, 172–193. https://www.jstage.jst.go.jp/article/gbkk/15/0/15_172/_article/-char/ja
- Suzuki, M., & Fujimori, H. (2014). Can-do risuto kaihatsu puroseshu ni okeru gakushū-sha no jiko hyōka to sono bunseki. *Tōkyōgaikokugodaigaku Ryūgakusei Nihongo Kyōiku Sentā Ronshū*, 40, 53–68.

- The Japan Foundation. (2017). *JF Nihongo sutandaado riyousha no tame no gaidobukeku (Shinpan)*. JF Standard for Japanese-Language Education. <https://jfstandard.jp/publicdata/ja/render.do>
- Tiantong, M., & Tongchin, P. (2013). A multiple intelligences supported web-based collaborative learning model using Stufflebeam's CIPP evaluation model. *International Journal of Humanities and Social Science*, 3(7), 157–165. <http://www.ijhssnet.com/journal/index/1735>
- Ting, Y.-L., Tai, Y., & Chen, J.-H. (2017). Transformed telepresence and its association with learning in computer-supported collaborative learning: A case study in English learning and its evaluation. *Interactive Learning Environments*, 25(3), 382–396. <https://doi.org/10.1080/10494820.2015.1131169>
- Universitas Negeri Semarang. (2008). *Panduan akademik [Academic guideline]*. Universitas Negeri Semarang.
- Yanagisawa, E. (2018). Ingurisshu track Nihongo puroguramu ni okeru reberu datōsei noo kenshō – Nihongo nōryokushiken mogi shiken to Can - Do statements chōsa no kekka kara. *Meijidaigaku Kokusai Nihon-Gaku Kenkyū*, 8(1), 93–118.
- Yanagisawa, E. (2019). Can-do statements o mochiita gakkai kaishi-ji to shūryō-ji ni okeru Nihongo nōryoku no jiko hyōka no hikaku - ET Nihongo ni okeru kyōiku no shitsu no hoshō ni mukete. *Meijidaigaku Kokusai Nihon-Gaku Kenkyū*, 11(1), 123–150.
- Yokota, T. (2014). A survey of non-native Japanese language teachers' viewpoints in Japanese A survey of non-native Japanese language teachers' viewpoints in Japan. *CAJLE Annual Conference Proceeding*, 239–247. https://www.cajle.info/wp-content/uploads/2014/09/Yokota_CAJLE2014_Proceedings_239-247.pdf