

## STRENGTHENING TEACHER COMPETENCE IN ANALYSING QUESTION ITEMS AT SDIT DARUL ABROR, GARUT REGENCY

Dian Rahadian<sup>1</sup>, Ayu Puji Rahayu<sup>2</sup>, Iman Nasrulloh<sup>3</sup>, Indra Budiman<sup>4</sup>

<sup>1,2,3,4</sup>Institut Pendidikan Indonesia, Garut, Indonesia

Email: ayupujirahayu14@institutpendidikan.ac.id

### Abstract

*Learning assessment is an essential component of the educational process as it evaluates students' achievement of competencies and provides a basis for improving instructional practices. However, evidence from the field shows that elementary school teachers still face challenges in conducting item analysis using quantitative approaches. This community service program was designed in the form of a workshop to enhance teachers' conceptual understanding of learning assessment as well as their technical skills in item analysis using Anates software. The program was implemented through pre-test, lectures, hands-on practice, group discussions, and post-test activities involving 52 teachers, with 41 serving as research respondents. Data analysis using the Wilcoxon test indicated a significant difference between participants' abilities before and after the workshop, with greater improvement observed in technical item analysis skills compared to conceptual understanding. These findings demonstrate that practice-oriented training supported by technology can effectively strengthen teachers' competence in developing valid and reliable assessment instruments. The implications of this program extend beyond the partner school, offering a sustainable collaboration model between universities and elementary schools to foster teacher professional development.*

**Keywords:** Learning assessment; item analysis; Anates; workshop; elementary school teachers

### INTRODUCTION

Evaluation of learning outcomes is a crucial component of the learning process, aiming to objectively and measurably assess the extent to which students' competencies have been achieved. When conducted systematically and following scientific principles, evaluation can provide accurate information regarding the effectiveness of learning, the validity of test items, and the areas that require improvement in the learning process. The findings from this evaluation can serve as valuable feedback for teachers to improve and refine their learning planning and implementation (Ropii & Fahrurrozi, 2017). In the context of elementary education, teachers play a crucial role in designing and analysing evaluation tools, particularly multiple-choice questions, which are still widely used in both formative and summative assessments.

Nevertheless, several studies have revealed that teachers' ability to analyse test items remains relatively low. In practice, teachers often construct test items based on estimates or intuition, without conducting statistical measurements of validity, reliability, difficulty level, discriminating power, or effectiveness of distractors (Putri & Abimanyu, 2024; Yaqin, 2023). However, item analysis is a crucial step in ensuring that evaluation instruments are truly capable of measuring the targeted competencies (Mawardi et al., 2023; Musyafa et al., 2023).

At SDIT Darul Abror in Garut Regency, the community service team, through a preliminary study using observations and unstructured interviews, found that although some teachers were accustomed to developing questions, their ability to analyse the results quantitatively and in a structured manner was still lacking. This condition can impact the quality of learning evaluation and impact educational decision-making, including determining remedial needs, enrichment programs, and final summative assessment results. Similar studies by Amelia et al. (2021) and Aquatika et al. (2022) have shown that training in the use of software such as Anates can help improve teachers' understanding of the concepts of question validity and reliability, as well as enhance the accuracy of learning outcome evaluation.

Based on these issues, this community service activity was designed as a workshop integrating two main focuses: (1) a comprehensive understanding of the concept of learning outcome evaluation, and (2) technical training in item analysis using Anates software. This approach was chosen because it is practical, aligned with teachers' needs in the field, and can have a direct impact on improving the quality of classroom learning. The use of Anates as an analysis tool has proven effective in helping teachers obtain important statistical data from the questions they develop (Augustia et al., 2025; Fiska et al., 2021).

Based on the background above, the research questions to be addressed through this Community Service (PKM) activity are: 1). How can teachers improve their understanding of the concept of scientific and systematic learning outcome evaluation?; 2). How can teachers improve their technical skills in analysing test items using Anates software?; 3). Is there a significant difference in workshop participants' understanding of learning outcome evaluation and item analysis using Anates software before and after the activity?

Furthermore, based on the research questions above, the objectives of this community service activity are as follows: 1). To improve teachers' understanding of the concept of scientific and systematic learning outcome evaluation; 2). To improve teachers' technical skills in analysing test items using Anates software. 3). To analyse whether there is a difference in participants' understanding of learning outcome evaluation and item analysis using Anates software before and after the activity.

This community service activity is designed to have a tangible impact. It is expected to provide direct benefits, particularly in increasing teachers' capacity to understand the principles of learning outcome evaluation and master the technique of item analysis using Anates software. With this understanding, teachers can develop more accurate and data-based assessment instruments, so that decisions regarding remediation, enrichment, and learning improvements are no longer based solely on estimates.

Furthermore, the implementation of the training results contributes to improving the quality of classroom learning and encourages schools to develop a more systematic evaluation culture. This program also serves as a means of collaboration between universities and schools to strengthen teacher professionalism and enrich academics' practical experience in carrying out their community service functions.

Problem-solving is carried out through a participatory and hands-on workshop approach. The workshop activities are designed interactively using lectures, discussions, case studies, and hands-on practice using Anates software. Direct, expository learning support emphasises the verbal delivery of material by tutors, effectively imparting the desired specific skills. The material is presented in stages, starting with the basics of learning evaluation, the concept of item analysis, and the practice of analysing questions based on student data using Anates software. Each session is led by competent speakers who hold doctoral degrees in education and specialise in educational technology and instructional design. It is hoped that the presence of these experts will contribute to improving the quality of learning outcome evaluation at SDIT Darul Abror, Garut Regency. Additionally, with the active involvement of teachers during the workshop, both theoretical understanding and practical skills can be mastered in a balanced and sustainable manner (Cahyaningrum et al., 2025; Härmälä, 2023).

On the other hand, this activity has novelty because it integrates three approaches rarely implemented comprehensively at the elementary school level: conceptual training on learning outcome evaluation, reinforcement of item analysis concepts, and software-based technical training. Unlike conventional training, which focuses on only one aspect, this program targets the simultaneous integration of conceptual and practical aspects. Furthermore, although the Anates software, based on classical test theory, has been proven effective in various studies (Amelia et al., 2021; Mawardi et al., 2023), its widespread use in elementary schools remains rare, particularly in areas such as Garut Regency. The following are some community service (PkM) projects that have been conducted in the form of workshops or training sessions on the topic of learning outcome evaluation. This data is taken from the Google Scholar database from 2021 to the present.

Aslam et al. (2021) conducted community service at the elementary school level to increase knowledge and deepen teachers' understanding of learning evaluation, particularly regarding assessment in the 2013 curriculum. The final results of this activity demonstrated an increase in teacher competence in conducting evaluations in accordance with the 2013 curriculum. Irawati et al. (2021) conducted community service with elementary school teacher working groups (KKG) in Sumedang Regency, aiming to increase teachers' understanding of developing evaluation tools using standardised instrument development procedures, as well as developing test instruments based on Higher Order Thinking Skills (HOTS). Meanwhile, Setiadi et al. (2021) conducted training aimed at enhancing the skills of Islamic high school (MA) teachers in developing instruments to measure students' attitudes and skills.

Furthermore, community service activities conducted by Ristiliana et al. (2022) aimed to provide information, knowledge, and skills to prospective teacher students regarding the effectiveness of using Anates in validating test instruments to support the implementation of learning evaluation in schools. Asriadi et al. (2022) focused on evaluating online learning for elementary school teachers, utilising workshops that included lectures, demonstrations, and practical exercises, as well as evaluating the training outcomes. Meanwhile, community service activities conducted by Prihatin et al.

(2022) focused on evaluating improved learning outcomes through the PLC training program at vocational high schools in the Surakarta region, utilising a scientific approach. These activities are expected to create superior human resources capable of competing and keeping pace with developments in industrial technology, particularly in the fields of mechatronics and PLC control systems.

Nurhijrah et al. (2023) conducted a community service aimed at broadening the horizons of teachers at MI Makassar in utilising technology-based learning media, specifically by using the Quizizz application for student learning evaluation. Furthermore, the community service conducted by Purba & Irwanoviana (2023) addressed the implementation of the Numbered Heads Together (NHT) learning model at Cinta Budaya Middle School in Medan City, to improve student achievement. Meanwhile, the community service conducted by Sudriyanto et al. (2023) focused on improving the quality of student learning outcome evaluation through the use of Google Drive and Google Forms applications in distance learning.

In 2024, Avana hosted a community service workshop on developing learning evaluation instruments in elementary schools. This workshop aimed to provide training related to the development of learning evaluation instruments, particularly at the elementary school level. Meanwhile, the community service conducted by Permatasari et al. (2024) focused on improving teachers' digital skills in designing creative and engaging evaluation. Salamah et al. (2024) addressed the topic of building a learning culture through learning communities to improve student learning outcomes. The method used in this community service was workshop training, which assisted in the formation and optimisation of learning communities in partner schools.

## RESEARCH METHODOLOGY

The community service method in this activity uses a participatory training approach and direct practice. At the same time, the research subjects in this workshop activity were 52 teachers of SDIT Darul Abror, Garut. The questionnaire was distributed randomly and voluntarily using *Google Form*, so that 41 people filled out the questionnaire from those present. The reasons for choosing the subject of SDIT Darul Abror elementary school teachers are as follows: first, SDIT Darul Abror is an elementary school located not too far from the campus, about 6 km to the east, which is expected to provide benefits to the surrounding environment. Second, SDIT Darul Abror is a developing elementary school that has young educators, who, of course, require guidance for the development of teacher human resource competencies. The workshop activity was held on June 17, 2025, in a room provided by the partner school. SDIT Darul Abror, Garut Regency, is located at the Putri Dinar Lestar Housing Complex, Tj. Kamuning, Tarogong Kaler District, Garut Regency. West Java.

The data collection tool used in this study was a questionnaire administered via Google Forms, both before and after the activity. This was intended to determine participants' understanding of the concept of learning evaluation and their item analysis skills, specifically using the Anates application. The questionnaire consisted of 24 items,

divided into two sections. The assessment scale used a Likert scale of 1-5. The first 12 items tested understanding of learning outcome evaluation, and the next 12 tested item analysis skills using Anates. Furthermore, the questionnaire included one open-ended question in the posttest, which aimed to determine key points learned during the activity and how to implement them in teaching practice. The research instruments are as follows.

**Tabel 1.** Pre-test and Post-test Instruments in Indonesian

No	A. Evaluation of Learning Outcomes
1	Saya memahami perbedaan antara pengukuran, penilaian, dan evaluasi.
2	Saya mengetahui tujuan utama dari evaluasi hasil belajar.
3	Saya memahami prinsip-prinsip evaluasi yang baik dan adil.
4	Saya dapat membedakan antara evaluasi formatif dan sumatif.
5	Saya tahu karakteristik evaluasi yang valid dan reliabel.
6	Saya memahami bahwa evaluasi harus sesuai dengan tujuan pembelajaran.
7	Saya mampu merancang instrumen evaluasi sesuai indikator pembelajaran.
8	Saya menyadari pentingnya analisis hasil evaluasi bagi perbaikan pembelajaran.
9	Saya mengetahui ragam teknik evaluasi (tes dan non-tes).
10	Saya mampu mengaitkan evaluasi hasil belajar dengan pencapaian kompetensi.
11	Saya memahami peran guru dalam proses evaluasi yang berkelanjutan.
12	Saya yakin mampu menyusun evaluasi yang mencerminkan capaian belajar siswa.
	<b>B. Item Analysis</b>
13	Saya mengetahui manfaat analisis butir soal terhadap kualitas asesmen.
14	Saya memahami apa yang dimaksud dengan tingkat kesukaran soal.
15	Saya memahami konsep daya pembeda dalam evaluasi soal.
16	Saya memahami fungsi pengecoh (distractor) dalam soal pilihan ganda.
17	Saya mengetahui langkah-langkah dasar menggunakan software ANATES.
18	Saya mampu menginput data siswa dan kunci jawaban ke dalam ANATES.
19	Saya dapat membaca dan menafsirkan hasil analisis dari ANATES.
20	Saya tahu bagaimana mengambil keputusan perbaikan soal berdasarkan hasil analisis.
21	Saya memahami kategori-kategori kualitas soal menurut hasil ANATES.
22	Saya mampu menyusun laporan sederhana dari hasil analisis butir soal.
23	Saya merasa percaya diri menggunakan software ANATES dalam praktik.
24	Saya termotivasi untuk menggunakan analisis soal sebagai bagian rutin dari evaluasi.
25	Apa hal paling penting yang Anda pelajari dari workshop ini, dan bagaimana Anda akan menerapkannya dalam kegiatan pembelajaran di kelas?

The research implementation procedures were as follows: 1). The activity began with preparing workshop materials and equipment; 2). A pre-test was conducted to determine initial understanding; 3). The activity continued with a presentation on learning outcome evaluation material, the concept of item analysis, and the stages of using the Anates software by two speakers; 4). Direct practice of question analysis using students' final semester exam results; 5). Group discussions and presentations of analysis results; 6). The activity concluded with a post-test and reflection.

After quantitative data collection using pre- and post-tests, the data analysis techniques used were descriptive and inferential statistical tests. Descriptive statistical tests were used to answer the first and second research questions, while inferential tests were used to answer the third research question. This statistical test also used a

comparison test of two correlated data groups. This test aimed to analyse whether there were differences in abilities before and after participating in the activity. This is important to determine the impact and effectiveness of community service activities on improving teachers' understanding or skills in evaluating learning outcomes and analysing item analysis.

## RESULTS AND DISCUSSION

This section presents the results of the data analysis conducted using IBM SPSS version 26. The analysis focuses on two main aspects: (1) descriptive statistics to identify the general trend of participants' scores before and after the workshop, and (2) inferential statistics to test whether the observed changes were statistically significant. The aim is to evaluate the effectiveness of the workshop in improving participants' ability to conduct learning outcome evaluation and item analysis through the use of Anates software. Based on the data analysis using IBM SPSS version 26, the following data were obtained.

### *Descriptive Analysis*

**Table 2.** Descriptive Analysis of Average Learning Outcome Evaluation Scores and Item Analysis

Respondents	Pre-test	Post-test	Pre-test	Post-test
	Evaluation of learning outcomes		Item analysis	
R1	3,7	3,0	3,0	3,0
R2	3,8	4,1	2,8	3,8
R3	3,4	4,0	2,6	4,0
R4	4,8	4,8	3,9	4,5
R5	4,1	4,7	3,3	4,9
R6	4,7	4,3	4,3	4,0
R7	4,1	4,8	3,7	3,8
R8	4,0	5,0	2,8	5,0
R9	4,0	4,3	2,6	4,1
R10	3,7	4,5	4,1	4,6
R11	3,7	4,8	2,9	4,8
R12	3,7	5,0	3,6	5,0
R13	3,8	3,8	3,4	3,3
R14	3,3	4,7	2,7	4,8
R15	3,9	5,0	2,8	4,9
R16	3,6	4,9	3,3	5,0
R17	3,2	4,0	3,1	4,0
R18	3,3	4,4	2,6	4,1
R19	3,7	4,6	3,3	4,5
R20	3,6	4,0	3,3	4,0
R21	3,7	4,3	3,6	4,3
R22	3,5	3,9	2,6	3,8
R23	2,8	4,0	2,7	4,0
R24	4,4	4,8	3,8	4,9
R25	3,4	1,9	3,0	1,9
R26	4,0	3,0	2,8	3,0
R27	4,0	4,0	2,9	4,0

R28	4,0	4,0	2,8	4,0
R29	3,6	4,0	2,9	3,3
R30	4,1	4,3	3,8	3,5
R31	4,3	4,0	4,1	4,0
R32	3,9	5,0	3,8	5,0
R33	3,7	4,5	3,3	4,8
R34	4,2	4,7	2,6	4,3
R35	4,0	4,9	2,6	5,0
R36	4,1	4,2	3,3	4,0
R37	3,1	5,0	2,8	4,3
R38	3,5	4,9	1,8	4,7
R39	4,0	4,7	2,9	5,0
R40	4,3	4,0	4,1	4,0
R41	3,0	3,6	2,5	3,5
<b>Average</b>	<b>3,8</b>	<b>4,3</b>	<b>3,1</b>	<b>4,2</b>

As presented in Table 2, the distribution of respondents' pretest and posttest scores shows a clear improvement across both variables, namely evaluation of learning outcomes and item analysis. The average pretest score for the learning outcome evaluation was 3.8, increasing to 4.3 in the posttest, which represents a 13% improvement. Meanwhile, the average score for item analysis increased from 3.1 to 4.2, indicating a more substantial improvement of 33%. These findings highlight that the workshop activities provided meaningful benefits, particularly in strengthening participants' analytical skills.

**Table 3.** Descriptive Analysis of Statistical Tests

		Statistics	
		Pre_test	Post_test
N	Valid	41	41
	Missing	0	0
Mean		3.37	4.32
Median		3.00	4.00
Mode		3	4
Std. Deviation		.536	.687
Sum		138	177

Table 3 provides a summary of descriptive statistics for the pretest and posttest results. The mean score increased significantly from 3.37 in the pretest to 4.32 in the posttest, reflecting an overall gain of 28%. Furthermore, the upward shift in both the median and mode values indicates that the majority of participants experienced consistent improvement. Although the standard deviation slightly increased from 0.536 to 0.687, this variation suggests that while most participants benefitted from the intervention, the degree of improvement varied among individuals.

### ***Inferential Analysis***

Inferential analysis is a statistical analysis commonly used to test research hypotheses. The research hypothesis (H1) in this study is "There is a significant difference in workshop participants' understanding of learning outcome evaluation and item analysis using Anates software before and after the activity." Therefore, the first step is to test the data for normality. If the data distribution is normal, the parametric paired t-test will be used. However, if the data distribution is not normal, the non-parametric Wilcoxon test will be used (Sugiyono, 2012; Sundayana, 2020).

**Table 4.** Pretest Normality Test

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.159	41	.011	.968	41	.287

a. Lilliefors Significance Correction

Table 4 reports the results of the pretest normality test using the Kolmogorov-Smirnov and Shapiro-Wilk methods. For the Shapiro-Wilk test, which is considered more accurate for samples below 50, the significance value obtained was 0.287. Since this value exceeds the threshold of 0.05, the pretest scores can be considered normally distributed. This result establishes an important basis for further inferential analysis, for this result ensuring the selection of appropriate statistical tests.

**Table 5.** Posttest Normality Test

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Posttest	.139	41	.045	.885	41	.001

a. Lilliefors Significance Correction

Table 5 presents the normality test results for the posttest scores. The Shapiro-Wilk significance value was 0.001, which is lower than the significance level of 0.05. This outcome indicates that the posttest data do not meet the assumption of normality. Therefore, because one of the normality test results was not normally distributed, the analysis was continued with a non-parametric test, specifically the Wilcoxon Signed Rank Test, to ensure the accuracy of the hypothesis testing.

**Table 6.** Wilcoxon Test

Related-Samples Wilcoxon Signed Rank Test

Summary

Total N	41
Test Statistic	799.000
Standard Error	77.148
Standardized Test Statistic	4.777
Asymptotic Sig.(2-sided test)	.000

Table 6 shows the outcome of the Wilcoxon Signed Rank Test, which was applied due to the non-normal distribution of the posttest data. The test result yielded a significance value of 0.000, which is far below the 0.05 threshold. This finding confirms that the observed improvement between the pretest and posttest scores is statistically significant. Therefore, it can be concluded that the workshop effectively enhanced participants' skills in evaluating learning outcomes and performing item analysis using Anates software.

Overall, the descriptive and inferential analyses consistently demonstrate that the workshop had a positive and statistically significant impact on participants' competencies. The improvements in both the evaluation of learning outcomes and item analysis scores indicate that the training using Anates software was effective in enhancing participants' understanding and practical skills, thereby providing a solid foundation for further discussion of its implications.

## **DISCUSSION**

### **1. Learning Outcome Evaluation**

In this section, the author will explain how teachers' understanding of the concept of evaluating scientific and systematic learning outcomes has improved. The results showed an increase in teachers' understanding of the concept of learning outcome evaluation after attending the workshop, as evidenced by an increase in the average pretest score from 3.8 to 4.3 on the posttest (a 13% increase). This increase suggests that the workshop enhanced teachers' understanding of scientific and systematic evaluation principles. These findings align with Arikunto's (2019) opinion that evaluations conducted in a structured manner will help teachers obtain an objective picture of student learning achievements and the effectiveness of the assessment instruments used.

Before the training, some teachers tended to develop questions based on experience and intuition without considering technical criteria, such as validity, reliability, and item discrimination. This was also found in a preliminary study at Darul Abror Islamic Elementary School, which showed that teachers' ability to analyse evaluation results quantitatively was still limited. After the workshop, teachers' conceptual understanding improved because they gained a theoretical foundation regarding the importance of data-driven evaluation. According to Mardapi (2017), learning outcome evaluation serves not only as a measure of student achievement but also as a basis for improving the curriculum, teaching methods, and assessment strategies.

Furthermore, this increased understanding strengthens teachers' position as evaluators capable of assessing the extent to which learning objectives are achieved. This is important because teachers' role is not only as learning facilitators but also as data-driven decision-makers. In line with the findings of Asi et al. (2024) systematic training

can help teachers understand the procedures for developing evaluation instruments according to established standards, thereby reducing their reliance on intuitive judgment.

The results of this study also support the views of Nitko and Brookhart, who emphasise that effective evaluation must be grounded in an understanding of the basic concepts of educational measurement (Brookhart & Nitko, 2019). This understanding enables teachers to develop fairer and more objective instruments, as well as to identify remedial and enrichment needs. In this context, the workshop successfully provided relevant and practical training, enabling teachers to integrate evaluation theory with the practice of question development and analysis.

Improving teachers' understanding of learning outcome evaluation also impacts the quality of classroom learning. When teachers can assess student competency using valid and reliable instruments, follow-up learning activities, such as remediation, enrichment, and the development of learning strategies — can be more targeted. This aligns with the findings of Aslam et al. (2021) Aslam et al., who showed that evaluation training for elementary school teachers contributed to improving their competency in managing assessments in accordance with curriculum requirements.

Therefore, the learning outcome evaluation workshop at SDIT Darul Abror had a positive impact on improving teachers' conceptual understanding. This strengthening provides a crucial foundation for teachers to develop a more systematic, scientific, and data-driven evaluation culture, ultimately enhancing the quality of learning in elementary schools.



**Figure 1.** Opening Ceremony

Figure 1 shows the opening ceremony, attended by a committee and a team of speakers from Institut Pendidikan Indonesia, Garut, and the Principal of SDIT Darul Abror Garut.



**Figure 2.** Presentation of the Concept of Learning Outcome Evaluation

Figure 2 presents a learning outcome evaluation by Dr. Dian Rahadian, M.Pd., an expert in educational technology and instructional design. Participants were attentively listening to the speaker's presentation.

## **2. Item Analysis Using Anates**

The results of the study showed that teachers' skills in analysing test items significantly improved after participating in a workshop on using the Anates software. The average pretest score of 3.1 increased to 4.2 on the posttest, representing a 33% increase. These findings demonstrate that hands-on training has a positive impact on teachers' understanding of the concepts of validity, reliability, difficulty level, discriminating power, and distractor effectiveness in multiple-choice test items.

Item analysis is a crucial component of learning evaluation, as valid and reliable instruments provide a more accurate picture of student competency achievement (Brookhart & Nitko, 2019). In practice, many teachers still rely on intuition when constructing test items, often resulting in uncertain instrument quality (Amelia et al., 2021). In fact, through quantitative analysis, teachers can ensure that test items are truly aligned with the learning objectives and differentiate between students who have mastered the material and those who have not.

The use of Anates has been proven to facilitate teachers' rapid, accurate, and data-driven analysis. Several previous studies support these findings. For example, Mawardi et al. found that training in the use of Anates improved teachers' competence in identifying invalid test items and improved the quality of evaluation instruments (Mawardi et al., 2023). Similarly, Ristiliana et al. reported that student teachers gained a better understanding of the concept of test analysis after using Anates as an instrument validation tool (Ristiliana et al., 2022).

Beyond its practicality, Anates also provides a technology-based learning experience that is relevant to current developments. Teachers not only acquire technical skills but also develop a scientific attitude in conducting evaluations, relying on empirical

evidence rather than mere guesswork (Aslam et al., 2021). Thus, the application of item analysis using Anates contributes to improving the overall quality of learning, as decisions made whether regarding remedial, enrichment, or final evaluation—are based on reliable data.

Conceptually, this activity also supports the strengthening of teachers' professional competence, as stipulated in the teacher competency standards. Teachers' ability to develop effective instruments and analyse learning outcomes is an important indicator in ensuring the quality of education in elementary schools (Irawati et al., 2021). Therefore, integrating learning outcome evaluation theory with software-assisted analysis practices, such as Anates, can serve as an effective training model to improve teacher professionalism, particularly in areas with limited access to technology-based training.

Therefore, the use of Anates in this workshop not only improved teachers' technical skills in item analysis but also strengthened their conceptual understanding of the importance of data-driven evaluation. This aligns with efforts to build a culture of systematic evaluation in elementary schools and contributes to the continuous improvement of learning quality.



**Figure 3.** Presentation of Item Analysis

Figure 3 presents an item analysis by Dr. Ayu Puji Rahayu, M.Pd., an expert in educational technology, curriculum, and teaching methodology. Participants enthusiastically listened to the material on item analysis so that it could be implemented in classroom learning practices.



**Figure 4.** Hands-on Practical Activity: Analysing Question Items Using the Anates Application

Figure 4 shows participants directly practising the use of the Anates application to analyse final exam results. This activity coincided with the end of the End-of-Semester Assessment (PAS) at the elementary school level in Garut Regency. Therefore, this activity significantly improved lecturers' competency in understanding the concepts of evaluation and analysis of student learning outcomes.

### **3. Differences in Workshop Participants' Understanding of Learning Outcome Evaluation and Item Analysis Using Anates Software Before and After the Community Service Activity**

Data analysis reveals that the workshop had a significant impact on teacher competency, improving both conceptual understanding and technical skills. The Wilcoxon Signed-Rank Test results showed a significance value of 0.000 ( $<0.05$ ), indicating a significant difference between pretest and posttest scores for both variables studied. Therefore, this training program was effective in developing teacher capacity in both learning outcome evaluation and item analysis.

Upon closer examination, teachers' understanding of the concept of learning outcome evaluation increased from an average score of 3.8 in the pretest to 4.3 after the posttest, representing an increase of approximately 13%. This improvement indicates that the participants understood the material presented on basic evaluation principles, including validity, reliability, discriminatory power, and instrument suitability for achieving learning objectives. However, improvements in the conceptual aspect were relatively moderate. This can be explained because mastering theory is an abstract concept that requires a longer process of internalisation through reflection and repeated application in the classroom context (Amelia et al., 2021; Brookhart & Nitko, 2019).

Meanwhile, teachers' skills in analysing test items using the Anates software showed more significant progress. The average score increased from 3.1 to 4.2, representing an approximate 33% increase. This increase suggests that the use of hands-on learning strategies leads to more rapid improvement in teacher skills. Anates facilitates teachers in calculating important indicators such as difficulty level, discriminating power, reliability, and distractor effectiveness more efficiently. This finding aligns with research by Ristiliana et al. and Mawardi et al., which states that using Anates can strengthen the practical competence of teachers and prospective teachers in assessing test quality (Mawardi et al., 2023; Ristiliana et al., 2022).

The difference in the magnitude of improvement between conceptual understanding and technical skills confirms that these two aspects have distinct characteristics. Technical skills develop more quickly because they are concrete and applicable, while conceptual understanding takes longer to internalise fully (Latifah et al., 2023). Therefore, the integration of theory and practice is crucial: conceptual

understanding provides a scientific framework for evaluation, while technical skills ensure the implementation of valid, reliable, and data-driven instruments.

These results also align with the findings of Amelia *et al.* (2021), who demonstrated that item analysis training assisted by Anates improved elementary school teachers' understanding of the concepts of instrument validity and reliability. Furthermore, Nurhijrah *et al.* (2023) expanded teachers' insights at MI AL ABRAR Makassar in utilising technology-based learning media through a community service program in Makassar, specifically by using the Quizizz application in the learning evaluation process for students. Then, the community service activities of Ristiliana *et al.* (2022) demonstrated that the use of Anates software in validating test instruments significantly strengthened student teachers' understanding of the concept of item analysis, thereby enriching authentic evaluation-based learning practices.

Overall, the results of this study confirm that workshops with an integrative approach can have a dual impact: strengthening teachers' conceptual foundations for evaluating learning outcomes while simultaneously equipping them with practical skills in item analysis. Both complement each other, enabling teachers to make more accurate learning decisions and foster a culture of systematic evaluation in elementary schools, particularly at SDIT Darul Abror Garut.



**Figure 5.** Group Photo After the Event

Figure 5 shows a group photo at the end of the event, featuring the committee, resource persons, and participants, including teachers at Darul Abror Islamic Elementary School (SDIT Darul Abror). This activity is expected to continue as a form of synergy between the Institut Pendidikan Indonesia (IPI) Garut and the SDIT Darul Abror educational unit, located near the campus. This will ensure the university's presence can make a tangible contribution to improving the quality of education, particularly in Garut Regency, West Java.

Analysis of one open-ended response reflecting on this community service activity indicates that the learning outcome evaluation workshop had a positive impact on teachers' understanding and skills. Many participants expressed that they now have a

clearer understanding of the concept of evaluation, including the distinction between measurement, assessment, and evaluation, as well as the importance of item analysis. Teachers also felt more skilled in identifying question characteristics, such as difficulty level, discriminating power, and effectiveness of distractors. This understanding is considered beneficial for improving the quality of instruments used in the learning process. Thus, the workshop not only strengthened theoretical aspects but also provided practical skills that can be directly applied in the classroom context.

Furthermore, almost all respondents highlighted the usefulness of technology, particularly the ANATES application, in facilitating the item analysis process. They considered using this application faster, more accurate, and more efficient than manual methods. Several teachers also emphasised that ANATES made it easier for them to compile simple yet informative analysis reports, which could then be used to improve subsequent evaluation instruments. The implication of these findings is a new awareness among teachers to routinely implement item analysis, thereby continuously improving the quality of learning evaluations.

On the other hand, although this activity demonstrated significant improvements in teachers' understanding of learning outcome evaluation and item analysis skills using Anates, several points are worth noting for reflection. First, the relatively short workshop time limited participants' opportunity to explore the material in depth and engage in repeated practice. Second, the research scope was limited to a single partner school, making the results contextual and not fully generalisable to schools with different characteristics. In this regard, further research or community service is recommended to involve participants from various schools, extend the duration of training to maximise skills deepening, and add follow-up evaluations to assess the sustainability of teacher skills implementation in the classroom.

## **CONCLUSION**

The community service activity, which included a workshop on learning outcome evaluation and item analysis using Anates software at Darul Abror Islamic Elementary School in Garut Regency, proved effective in improving teacher capacity. The Wilcoxon test results showed a significant difference between teachers' understanding and skills before and after the training, with greater improvement in item analysis than in conceptual understanding. These findings confirm that a hands-on approach, supported by statistical software, can enhance teachers' competencies in designing valid, reliable, and data-driven assessment instruments.

However, the effectiveness of the activity still needs to be improved through broader and more sustainable implementation. Developing a similar program on a larger scale, involving multiple schools, and extending the training duration is expected to

strengthen the impact and sustainability of the results. Thus, this activity not only contributes to improving the quality of learning in partner schools but also serves as a model for implementing collaboration between universities and elementary schools in developing teacher professionalism in the field of learning outcome evaluation.

### **Acknowledgments:**

The author would like to thank the Institut Pendidikan Indonesia (IPI) Garut, for funding this community service project under contract number 048/IPI.L3/Kontrak.PkM/VIII/2025

### **DAFTAR PUSTAKA**

- Amelia, N., Paridjo, & Sina, I. (2021). Analisis butir soal penilaian akhir tahun mata pelajaran matematika menggunakan software anates. *AKSIOMA: Jurnal Matematika Dan Pendidikan Matematika*, 12(3), 447–456.
- Aquatika, F., Rofi'i, & Pratama, M. H. B. (2022). Penggunaan Aplikasi Anates pada Analisis Butir Soal Penilaian Tengah Semester Bahasa Indonesia Kelas X MIPA 8 SMA Muhammadiyah 2 Surabaya. *Jurnal Metalanguage*, 4(4), 1–7.
- Arikunto, S. (2019). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Asi, N., Fauzi, I., & Nugraha, R. F. (2024). Pelatihan Penyusunan Instrumen Evaluasi Berbasis HOTS Melalui Model Literasi Numerasi Bagi Guru Bahasa Inggris SMP di Kota Palangka Raya. *Jurnal Abdimas Prakasa Dakara*, 4(1), 13–24. <https://doi.org/10.37640/japd.v4i1.1884>
- Aslam, Handayani, S. L., & Zulherman. (2021). Peningkatan Kemampuan Guru SDN Ngargogondo Magelang dalam Melakukan Penilaian Melalui Workshop Evaluasi Kurikulum 2013. *Jurnal Masyarakat Mandiri*, 5(5), 2195–2201. <https://doi.org/https://doi.org/10.31764/jmm.v5i5.5337>
- Asriadi, A., Achmad Shabir, & Muhammad Farid. (2022). Workshop Evaluasi Pembelajaran Daring bagi Guru Sekolah Dasar Mitra PGSD UNM Kampus VI Bone. *WAHATUL MUJTAMA': Jurnal Pengabdian Masyarakat*, 3(1), 1–10. <https://doi.org/10.36701/wahatul.v3i1.451>
- Augustia, A. D., Agustia, C. N., Azzahra, D., Hakim, L., & Pratiwi, V. (2025). Analisis Validitas dan Reliabilitas Soal Pilihan Ganda dengan Menggunakan Software Anates pada Mata Pelajaran Perpajakan. *Analisis Validitas Dan Reliabilitas Soal Pilihan Ganda Dengan Menggunakan Software Anates Pada Mata Pelajaran Perpajakan*, 5(1), 250–265.
- Avana, N. (2024). Workshop Pengembangan Instrumen Evaluasi Pembelajaran Di Sekolah Dasar. *Jurnal Pengabdian Pendidikan Masyarakat (JPPM)*, 5(1), 179–183. <https://doi.org/10.52060/jppm.v5i1.1425>
- Brookhart, S. M., & Nitko, A. J. (2019). Educational Assessment of Students. In *Pearson Education, Inc., Upper Saddle River, NJ*. Pearson.
- Cahyaningrum, D., Hasany, T. D., Asri, Ni Wayan Adelia Mutiara Safitri, D., & Ibrahim, I. D. K. (2025). Pelatihan Strategi Digital Branding dalam Upaya Optimalisasi Bank Sampah. *Jurnal Pengabdian Kepada Masyarakat Nusantara (JPkMN)*, 6(1), 196–205. <https://doi.org/http://doi.org/10.55338/jpkmn.v6i2.5558>

- Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis butir soal ulangan harian ipa menggunakan software anates pada pendekatan teori tes klasik. *Jurnal Natural Science Educational Research*, 4(1), 65–76.
- Härmälä, M. (2023). Research Quarterly Finnish Compulsory School English and Swedish Teachers' Classroom Assessment Practices. *Language Teaching Research Quarterly*, 37, 108–124. <https://doi.org/https://doi.org/10.32038/ltrq.2023.37.05>
- Irawati, R., Kurnia, D., Nugraha, D., & Sunarya, D. T. (2021). Pelatihan Perancangan Alat Evaluasi Untuk Menilai Kompetensi Siswa Sekolah Dasar Di Kecamatan Surian Kabupaten Sumedang. *Jurnal Pasca Dharma Pengabdian*, 2(2), 91–95. <https://ejournal.upi.edu/index.php/crecs/article/view/14284>
- Latifah, A., Husaini, F., & Khoirotun Nisa, A. (2023). Pengembangan Instrumen Penilaian Berbasis Hots. *Didaktik : Jurnal Ilmiah PGSD STKIP Subang*, 9(2), 4486–4496. <https://doi.org/10.36989/didaktik.v9i2.1057>
- Mardapi, D. (2017). *Pengukuran, penilaian, dan evaluasi pendidikan*. Mitra Cendekia.
- Mawardi, M. S., Fuady, A., & Sunismi. (2023). Analisis Butir Soal Pilihan Ganda Menggunakan Anates pada Penilaian Tengah Semester Kelas VII D SMP Negeri 1 Ngajum Kabupaten Malang. *Wahana : Tridarma Perguruan Tinggi*, 75(1), 31–41.
- Musyafa, A., Nasir, M., Jannah, F., & Sudadi. (2023). Implementasi Evaluasi Hasil Belajar Pendidikan Agama Islam Berbasis Kurikulum Merdeka. *Jurnal Pendas Mahakam*, 8(2), 209–215.
- Nurhijrah, Qur'ani1, B., Suryana, S., Hamsar, I., & Gawarti. (2023). Workshop Aplikasi Quizizz untuk Evaluasi Hasil Belajar di MI Al Abrar Makassar. *JEPKM: Jurnal Edukasi Dan Pengabdian Kepada Masyarakat* *Jurnal Edukasi Dan Pengabdian Kepada Masyarakat*, 2(1), 26–32.
- Permatasari, D., Dayu, K., & Setyaningsih, N. D. (2024). Pemberdayaan Guru Melalui Workshop Quizizz untuk Meningkatkan Keterampilan Digital pada Implementasi Kurikulum Merdeka. *Jurnal Pengabdian Masyarakat Bangsa*, 2(10), 4391–4396.
- Prihatin, J. Y., Pambudi, S., & Rahmat, R. (2022). Evaluasi Peningkatan Hasil Belajar Program P3M Pelatihan Plc Pada Smk Di Wilayah Surakarta Dengan Pendekatan Sainifik. *Jurnal Abdi Masya*, 2(1), 1–7. <https://doi.org/10.52561/abma.v2i1.203>
- Purba, F. O., & Irwanoviana, N. (2023). Penerapan Model Pembelajaran Numbered Heads Together Dalam Meningkatkan Hasil Belajar Siswa SMP Cinta Budaya/Chong Wen. *Jurnal Hasil Pengabdian Masyarakat (JURIBMAS)*, 2(2), 200–205. <https://doi.org/10.62712/juribmas.v2i2.127>
- Putri, J. A. Y., & Abimanyu, A. (2024). Pentingnya Evaluasi Hasil Belajar Di SMP. *Jurnal Ilmiah Kanderang Tingang*, 15(2), 422–432. <https://doi.org/https://doi.org/10.37304/jikt.v15i2.369>
- Ristiliana, Wati, I., Novita, Y., & Lubis, M. I. (2022). Pelatihan Penggunaan Software Anatest dalam Menganalisis Butir Soal Tes Ekonomi bagi Mahasiswa Calon Guru. *Tasnim Journal for Community Service*, 3(2), 63–69.
- Ropii, M., & Fahrurrozi, M. (2017). *Evaluasi Hasil Belajar*. Universitas Hamzanwadi Press.
- Salamah, E. R., Tiyas Rifayanti, Z. E., Trisnawaty, W., & Fitra Raharja, H. (2024). Membangun Budaya Belajar Melalui Komunitas Belajar Dalam Peningkatan Hasil Belajar Peserta Didik. *ABIDUMASY Jurnal Pengabdian Kepada Masyarakat*, 5(01), 37–43. <https://doi.org/10.33752/abidumasy.v5i01.5894>

- Setiadi, D., Wilian, S., & Sridana, I. N. (2021). Peningkatan Keterampilan Guru Ma Al-Aziziyah Puteri Kapek, Gunungsari Dalam Mengembangkan Instrumen Asesmen Sikap Dan Keterampilan Hasil Belajar. *Jurnal Interaktif: Warta Pengabdian Pendidikan*, 1(1), 20–25. <https://doi.org/10.29303/interaktif.v1i1.4>
- Sudriyanto Sudriyanto, Novi Fitriani, Dinda Lailatul Qomariyah, & Wiwik Widianti. (2023). Pelatihan Pemanfaatan Aplikasi Google Drive dan Google Form Sebagai Instrumen Evaluasi Hasil Belajar Siswa bagi Guru Madrasah Ibtidaiyah Raudlatul Mutaallimin. *ASPIRASI : Publikasi Hasil Pengabdian Dan Kegiatan Masyarakat*, 1(6), 145–155. <https://doi.org/10.61132/aspirasi.v1i6.89>
- Sugiyono. (2012). *Metode Penelitian Pendidikan (Ke-14)*. Alfabeta.
- Sundayana, R. (2020). *Statistika Penelitian Pendidikan (Kelima)*. Alfabeta.
- Yaqin, H. A. (2023). Evaluasi Hasil Belajar Pendidikan Islam. *JoeL: Journal of Educational and Language Research* *Journal of Educational and Language Research*, 3(2), 61–72.