The Development of CEFR-Based Nahwu and Shorof Learning Evaluation in *Madrasah Aliyah* in Cirebon

Zainal Arifin Ahmad[a,1], Irfan Hania[a,2], Faizmailiatus Sofa[a,3], Sri Utami[a,4], Putri Nurinadia[a,5]

[a]Sunan Kalijaga Islamic State University, Yogyakarta

[1]zainal.a@uin-suka.ac.id, [2]irfanhania0@gmail.com, [3]faizsofa1@gmail.com, [4]sriutami199712@gmail.com, [5]pn1632965@gmail.com

ARTICLE INFO	ABSTRACT
	This study aims to determine the implementation and development of CEFR-based nahwu and shorof learning evaluations at MA An-Nur Jagasatru Cirebon. The research methodology used is the research and development method. In the development of this research, we used the Rowntree model, which was consisted of three stages, namely the planning stage, the development stage, and the evaluation stage. The trial phase began with the validation of material experts and linguists. Then, after being revised, the product was tested on class X students, namely 8 students for testing prototype 1 and 21 students for testing prototype 2. The results showed that the average score of material experts was 85%, which meant the product was considered very good, and linguists 88%, which meant the product was considered very good. Then at the testing stage of prototype 1, an average score of 64.37, which meant the product was considered quite good, and at the testing stage of prototype 2, an average score of 74.71, which meant it was considered very good. Based on interviews with teachers, they obtained very good results. Based on the results of the study, it can be concluded that the evaluation of nahwu and shorof learning based on CEFR is feasible and effective as an evaluation tool in nahwu and shorof learning, especially in MA An-Nur Jagasatru Cirebon.
Keyword	CEFR; Development Learning Evaluation; Nahwu; Shorof;

مستخلص البحث

تهدف هذه الدراسة إلى تقويم تقييمات التعلم النحو و الصرف المستندة إلى CEFR في المدرسة الثانوية الإسلامية النور جاكاسطرو شربون. منهج البحث المستخدم هو طريقة البحث والتطوير. يتم تطوير هذا البحث باستخدام نموذج راونتري (*Rowntree*) الذي يتكون من ثلاث مراحل وهي مرحلة التخطيط، ومرحلة التطوير، ومرحلة التقييم. تبدأ مرحلة التجريبية بالتحقق من صحة خبير بالمواد واللغويين. ثم بعد المراجعة، تم التجريبية المنتج على طلاب الفصل X، وهم ٨ طلاب لتجريب النموذج المبدئي ١ و ٢١ طالبًا لتجريب النموذج المبدئي ٢. وأظهرت النتائج متوسط درجات ٨٥٪ من خبير المواد، مما يعني أن المنتج يعتبر جيدًا جدًا، واللغويون ٨٨٪، مما يعني أن المنتج تم تقييمه جيدًا. ثم في مرحلة التجريبية النموذج المبدئي ١، كان متوسط الدرجة ٦٤.٣٧، مما يعني أن المنتج يعتبر جيدًا جدًا، وفي مرحلة التجريبية للنموذج المبدئي ٢، بمتوسط درجة ٧١.٧٤، مما يعني أنه يعتبر جيدًا جدًا. بناءً على المقابلات مع المعلمين، استنادًا إلى نتائج الدراسة ، يمكن الاستنتاج أن تقييم تعلم النحو و الصرف على أساس CEFR ممكن وفعال كأداة تقييم في تعلم النحو و الصرف، خاصة في في المدرسة الثانوية الإسلامية النور جاكاسطرو شربون.

كلمات أساسية	CEFR؛ تطوير تقييم التعلم؛النحو؛الصرف

INTRODUCTION

Knowledge is included in a realm that develops over time (Ngafifi, 2014). The progress that occurs in the world makes all aspects of life become constantly changing with its development (Ali, 2021). This also applies to technology and science (Setiawan, 2018). The existence of advances that exist in technology and science makes developments that exist in the educational aspect (Jamun, 2018). Education, which is the main actor in educating students (Hermanto, 2020) has an important task in advancing knowledge (Nurkholis, 1970) by adjusting to evolving and changing needs (Kadi & Awwaliyah, 2017). Therefore, it is hoped that education can choose modern innovations and creations that can keep up with the times (Sibagariang et al., 2021).

The changes that occur in education should not only be in one aspect but include all aspects of education (Basmar et al., 2021). If development is carried out only in one aspect, it will lead to an imbalance in the world of education (Indayati et al., 2017). Therefore, it is recommended for developers to carry out periodic developments in each area of education, including the curriculum.

In learning activities, a component in the curriculum that participates in determining the success or failure of a learning activity is evaluation (Ardiansyah & Muhammad, 2020). Ralph Tyler defines evaluation in education as the activity of gathering information in order to measure and assess how, to what extent, and to what extent educational goals have been met (Ridho, 2018). Meanwhile, Cornbach and Stufflebeam suggested evaluation with measurements that were useful in determining further decisions (Arikunto, 2021). Since there is a development in science, the learning objectives also undergo a transformation (Lubis, 2016). And when the learning objectives are changed, the learning evaluation should also experience development (Zazkia & Hamami, 2021). Therefore, the development of learning evaluation should be carried out periodically (Hasibuan, 2016) and adjusted to the desired learning objectives (Rahayu & Listiyadi, 2015).

Looking at the international level, the evaluation of shorof and nahwu science learning in Indonesia has not been carried out properly (Wahyono, 2019). This is evidenced by the many problems that occur in teaching and learning activities (Sa'adah, 2019). Therefore, we need international standards that are commonly used in language learning. One of the internationally recognized standards in language learning is CEFR (Uri & Aziz, 2020). Various countries have adapted the CEFR in using as a standard for learning foreign languages (Foley, 2019). In its implementation using the standard level that has been set (NATSIUK & BABII, 2019). The CEFR has successfully improved the skills of foreign

languages learned such as writing skills (Zou & Zhang, 2017), reading skills (Poupounaki-Lappa et al., 2021), speaking skills (Yoshitomi, 2021), and listening skills (He & Chen, 2017). The CEFR (Common European Framework of Reference for Language) is a framework for studying foreign language competencies that is initiated by Europe and is currently included in international standards International (Yin & Ahmad, 2021). Therefore, CEFR is commonly used as a mecca for learning language competence (Talqis Nurdianto & Noor Azizi bin Ismail, 2020). With the CEFR being used as a standard mecca for foreign language learning, this study intends to utilize the CEFR as material for developing learning evaluations.

Development in evaluation is important to carry out, because the changes that occur in education are mainly in learning objectives. Various research and evaluation developments have been carried out to support effective Arabic learning. There are some previous studies such as by Mella Hanifa Nugraha with the title Development of Multimedia-based Arabic Learning Instruments using Wondershare Quiz Creator (Nugraha, 2017) and by Hurul Jinani entitled Development of Istima' and Qira'ah Evaluation Media Using Hot Potato 6 in Arabic language learning in class VIII in Islamic Junior High School (Jinani, 2013). These evaluation developments only focus on optimizing learning media. Basically, in carrying out a development on the evaluation tool, appropriate standards are also needed as the main benchmark. Appropriate evaluation standards are needed to revise the unclear standards used. Therefore, as a result, the purpose of this research is to create an evaluation of nahwu and shorof learning based on the CEFR (Common European Framework of Reference for Languages), specifically in MA An-Nur Jagasatru Cirebon.

METHOD

This CEFR-based evaluation of nahwu and shorof learning is developed using a research and development (R&D) type of research. The development carried out is in the form of developing test instruments. The development model in this study is the Rowntree model which consists of three stages, namely the planning stage, the development stage, and the evaluation stage. According to Prawiradilaga, the Rowntree model is a product-oriented model that focuses on making educational materials (Westomi et al., 2018)

The planning stage needs analysis and formulation of learning objectives. Then at the development stage, which is about topic development, drafting, production of prototypes of one type of product that will be used for learning. At the evaluation stage, the researcher used Tessmer's formative evaluation model,

namely: (1) self evaluation; (2) expert review; (3) one-to-one evaluation; and (4) small group evaluation, because the product that will be produced in this research is a test instrument in the form of questions according to the nahwu and shorof material that adapts the CEFR model.

The study was conducted at MA An-Nur Jagasatru Cirebon by requiring the role of material experts and linguists during development. Learning evaluation was developed and implemented to students in the learning process in class X MA An - Nur Jagasatru Cirebon. The revised product was tested on class X students as research subjects , namely on 8 students for prototype test 1 and 21 students for prototype test 2.

RESULTS & DISCUSSION

According to the results of the study, the average score obtained at the expert trial stage is as follows. The product is rated very good by material experts with an average score of 85%, and very good by linguists with an average score of 88%. The average score for the first prototype testing stage is 64.37 which indicates that the product is considered quite good, and in the second prototype stage it is 74.71 which indicates that the product is considered very good. Being very good, in the test stage to students. The results of teacher interviews are very good. The development of CEFR-based nahwu and shorof learning evaluations for class X MA An-Nur Jagasatru Cirebon students can be said to be good, but it still needs improvement according to expert advice, student suggestions as objects, and suggestions from the teacher.

The first stage of the research was planning that begins with the results of the Needs Analysis. Dwiyogo stated that needs analysis is the first step of three important steps in research and development activities. Then, product development and finally product testing. This research and development are intended to assess the evaluation of learning from nahwu and shorof material in Arabic learning. This learning evaluation is developed and used for students in the learning process in class X MA An-Nur Jagasatru Cirebon. The first step taken in this research is planning, the second is the development stage and the third is the evaluation stage (Egok & Hajani, 2018). Researchers at this stage observed student characteristics, interviewed the MA teacher An-Nur Jagasatru Cirebon, studied the learning used, and analyzed the material. This is done to provide an overview of the content of the nahwu and shorof material being taught, as well as to create general and specific learning objectives.

At first, the research was carried out by observing class X while doing nahwu and shorof learning in class (Permatasari & Aliyah, 2021). From the observations that the researchers made, the atmosphere in the classroom was less effective, due to the lack of nahwu and shorof learning standards, so it was

necessary to evaluate with the appropriate achievement standards in order to achieve the learning objectives to be addressed.

Based on the problems faced by researchers in the field and through the results of interviews with teachers, there are still some difficulties encountered, such as in determining appropriate standards for nahwu and shorof learning (Utomo et al., 2021). Students tend to be less motivated because the method used does not attract interest and the delivery of material is too long, learning is monotonous, and questions and answers that even most students are unable to answer. This can be seen when researchers observe how many students pay attention to other things besides the learning process during the learning process.

From the results of the needs analysis, it can be concluded that a new concept of achievement standards is needed at MA An-Nur Jagasatru Cirebon, so that students are more enthusiastic in the learning process. The material is delivered using the lecture method, question and answer and exercises with assignments. In general, students of class X still have difficulty in understanding nahwu and shorof material. Based on various kinds of materials and problems of teachers and students, so it is necessary to evaluate learning with achievement standards based on CEFR to assist teachers in determining the concept of learning achievement standards and make it easier for students to understand the material well. Therefore, researchers try to develop an evaluation of CEFR-based nahwu and shorof learning at MA An-Nur Jagasatru Cirebon.

The second stage is the development step. There are three stages carried out by the researchers, namely, 1) topic development, 2) drafting, and 3) prototype production in this development step. First is self-assessment that refers to the evaluation prototype for nahwu and shorof learning which is evaluated by the researchers their self (*self-evaluation*) to reduce errors in the delivery of material to match learning achievements in learning evaluation, then it will be validated by *expert review* and assessed on *a one-to-one evaluation basis*. So that teachers can use the resulting learning evaluation as a guide in making better learning plans(Egok & Hajani, 2018).

At the topic development stage, it is done by making or compiling a grid of nahwu and shorof material questions with the theme of *aqsamul fi'il (mu'annats, mudzakkar), dzorof zaman* and *dzorof makan* (مذكر)الفعل أقسام ومؤنث، (ظرف، ظرف زمان، ومكان). The created material must be in accordance with the basic competencies (KD) and achievement standards at the CEFR. Next is the draft preparation stage, which is entering the stage of making nahwu and shorof material questions with adjusted themes. Then, after compiling the grid and making questions, the next step is to produce a prototype. The nahwu and shorof questions that have been made will be given to students and tested in the form of tests. In this case, the researchers will develop an evaluation of CEFR-based

nahwu and shorof learning with the theme *aqsamul fi'il (mu'annats, mudzakkar),* وظرف زمان، ظرف ،(ومؤنث مذكر) الفعل أقسام *dzorof zaman* and *dzorof makan* (مكان). The theme is a learning evaluation design that was originally made in the form of 30 questions consisting of multiple choice including nahwu and shorof material, and the suitability of basic competencies with level B1 achievement standards at CEFR. The overall result of this development stage is also called prototype I.

The third stage is the evaluation stage after completing the planning and development stage, which seeks to assess the validity and application of the resulting educational material. Formative evaluation is a procedure used to collect information about how well the learning evaluation material achieves predetermined objectives. (Toriqul Arif, 2019) Tessmer's formative test, which will be discussed below, can be used to assess the acceptability of items, materials, and designs. Validity tests and student feedback on the desired output are also carried out at the evaluation stage. Tessmer evaluation is used to guide evaluation at this level, which includes self evaluation, expert review, one to one evaluation, small group evaluation, and field test evaluation (Dessiana Irma Yuanita, Hamdi Akhsan, 2015).

Self-Evaluation

At this evaluation stage, an evaluation of the CEFR-based nahwu and shorof learning that has been developed and reviewed by researchers is carried out on the product. Researchers assess the information used to teach nahwu and shorof based on CEFR achievement standards here. Is the material in accordance with the CEFR achievement standards? Are the language rules appropriate? And whether the grid of questions and test questions is in accordance with the predetermined theme? (Widiastuti et al., 2019).

Expert Review

This stage aims to obtain a valid CEFR-based nahwu and shorof learning evaluation. The results of prototype I, which was developed through *self-evaluation*, were given to experts (*experts*) to be validated. (Widyastuti Akhmadan, 2017). Validators validate and evaluate products that have been created by the researchers. Validity tests are carried out on material validation and language validation. The validation results are given in the form of percentages on the validation sheet and will be used as a reference to revise the CEFR-based nahwu and shorof learning evaluation materials as prototype I.

Validation Results of Nahwu and Shorof Material Experts

A draft evaluation of model I learning was submitted for *expert review*. (Muhammad Anggi Prasetya, 2021) This study seeks to obtain valid standards for achieving learning evaluation. Overall, the results of material expert validation were obtained with a percentage of 85% of validators I and II as follows:

Table 1. Material Expert Validation Results

Aspects	Valuation		Percentage	Total Percentage
	Validator I	Validator II		
Content Quality	4	4	80%	83%
	4	4	80%	
	4	5	90%	
	4	4	80%	
Accuracy of Coverage	4	5	90%	85%
	5	4	90%	
	4	4	80%	
	4	4	80%	
Inquiry	4	4	80%	80%
	4	4	80%	
	4	4	80%	
	4	4	80%	
	4	3	70%	
	4	4	80%	
	5	4	90%	
Language	5	3	80%	90%
	5	3	80%	
	5	3	80%	
	5	4	90%	
	5	4	90%	
Sum	**97**	**78**	**85%**	

The Table 1 has four aspects that are assessed by relevant experts (*expert review*) in the field of material, each aspect is scored namely from validator I and validator II. The aspects written in the expert validation of nahwu and shorof materials are: 1) the quality of the content (83%), this aspect has been adapted to the material *aqsamul fi'il (mu'annats, mudzakkar), dzorof of the age* and *dzorof of eating* (أقسام الفعل (ومؤننم مذكر)، ظرف ،مكان وظرف زمان), 2) the accuracy of coverage (85%) with the competencies that students must learn, in this case the material is appropriate, 3) inquiry (80%), that is to say In the presentation of nahwu and shorof material there is an analysis of language, and the rules of language, (4) language (90%), in this stage the language used is Arabic.

Based on the description, this case based on expert reviews the material can be concluded from the data it is clear that the learning evaluation that the researcher developed is valid and can be used by the school.

Linguist Validation Results

Linguists also agree on the draft learning evaluation model I. Linguists provide comments and recommendations as follows: During practice, students should understand the results of the exercise and which parts still need revision. Previously, in this evaluation, students did not receive grades or scores for completing practice questions. To ensure complete student understanding, practice questions can be repeated. Below is a presentation from linguists with an overall percentage of 88% of validators I and II as follows:

Table 2. Linguist Validation Results

Aspects	Valuation		Percentage	Total Percentage
	Validator I	**Validator II**		
Plays	4	5	90%	90%
	4	5	90%	
	4	5	90%	
Communicative	4	5	90%	90%
Dialogical and Interactive	4	5	90%	90%
CEFR compliance	4	5	90%	90%
	4	5	90%	
Conformity with language rules	4	4	80%	85%
	4	4	80%	
	4	5	90%	
	4	5	90%	
Use of terms, symbiols, and icons	4	5	90%	83%
	4	4	80%	
	4	4	80%	
Sum	**56**	**66**	**88%**	

In addition, the draft learning evaluation model I is supported by the six indicators in the Table 2 that have been assessed by experts who have reviewed the language field. The percentage of written indicators contained in the evaluation of learning on the theme of *aqsamul fi'il (mu'annats, mudzakkar), dzorof zaman* and *dzorof makan* (أقسام الفعل (ومؤنث مذكر)، ظرف ،ظرف زمان ومكان), specifically as follows: 1) Straightforward for the purpose of being able to speak, in the aspect of language proficiency objectives must be in accordance with the material made, 2) Communicative, in this aspect the language in practice questions is appropriate, 3) Dialogical and Interactive, in this aspect the presentation of the language has in accordance with the material seen in the grid and practice questions, 4) CEFR compatibility, this aspect uses achievement

standards originating from Europe, namely CEFR is appropriate and makes it easier for students to understand the material well, 5) Compliance with language rules, this is already exist and can be seen from the components in the learning evaluation that have been adapted to the language rules in the material made, 6) The use of terms, symbols, and icons, can be seen from the existence of terms, symbols, and icons in questions that make it easier for students to understand the meaning of the material and the rules themselves.

One to One Evaluation

Then a small group trial will be conducted using a product that has been improved based on the validation results of material experts and linguists. The interview sheet is used as a data collection tool in (Egok & Hajani, 2018) a one to one evaluation trial. To get student feedback on the learning evaluation prototype with the theme *aqsamul fi'il (mu'annats, mudzakkar), dzorof zaman* dan *dzorof makan* (أقسام الفعل (مذكر ومؤنث)، ظرف زمان، وظرف مكان), interviews were conducted with students to find out their opinions.

To describe the target population of low, medium, and high-ability children, the researchers will choose eight children at random. According to pre-research data, high ability means they can complete sentences, translate a sentence into Arabic or vice versa, and sort imperfect sentences based on their shape. Students with moderate abilities can recognize and complete sentences in nahwu and shorof materials, while students with low abilities can only recognize the basic forms of nahwu and shorof. Prototype 1 will be used in the material presented to students. Students will be observed again during the learning process using the observation sheet that has been provided to identify and evaluate problems that arise during prototyping 1 evaluation of CEFR-based nahwu and shorof learning. Prototype 2 of the learning evaluation, which has undergone expert testing, is used for small group evaluation after prototype 1 is proven to be feasible in a one-to-one evaluation.

Table 3. Prototype 1 Results

Prototype 1		Number of Samples
Average Values	64.37	
Minimum Value	25	8
Maximum Value	87	

According to the results of one-to-one evaluation and direct observation with several students in Table 3, they find it easy to understand the material using the CEFR achievement standards because their basic competencies are in accordance with the existing level in CEFR and they are able to understand the

rules of nahwu and shorof well. Students are more motivated to learn about each subject matter when they are involved in structured and interesting learning activities.

Small Group Evaluation

The evaluation of basic nahwu and shorof learning was tested in small groups after prototype I was revised based on the results of an expert review of the problems faced by teachers and students during testing on the first prototype (one-to-one evaluation). 26 students of MA An-Nur Jagasatru Cirebon who did not take part in the one-to-one evaluation were given an explanation of the revision of *aqsamul fi'il (mu'annats, mudzakkar), dzorof zaman* and *dzorof makan* (أقسام الفعل مذكر ومؤنث، ظرف، ظرف زمان وظرف مكان) (second prototype).

Currently, 21 students will be divided into small groups to test the results of the prototype 2. The second revised prototype was then taught in front of the class. The use of nahwu and shorof learning assessments (prototype 2) by students in the ongoing learning process will be directly monitored and evaluated as long as they are still learning. This is an attempt to determine how significant the practical component of learning evaluation is from the learner's perspective. The product has been completed after the necessary improvements have been made and prototype 2 has proven valid or useful. (Widiastuti et al., 2019)

Table 4. Prototype 2 Results

Prototype 2		Number of Samples
Average Values	74.71	
Minimum Value	53	21
Maximum Value	89	

Based on the results of the research in Table 4, the development of nahwu and shorof evaluations in this study has theoretical and practical implications. Theoretical implications can support previous studies related to the evaluation of nahwu and shorof learning. Like the research conducted by (Chaqoqo, 2017) and regarding the evaluation of nahwu learning, the practical implications are shown by the empirical evidence obtained in this study. The results of this study can be used as a basic reference when CEFR learning standards are used in Indonesia.

Basically, learning evaluation is a process and not a result. How is the process of giving consideration to the quality of learning. In reality, the evaluation aims to find the quality of something. In learning evaluation activities, the first step that must be done by the teacher is to pay attention to the purpose of the evaluation. There are general and specific evaluation objectives. Teachers can use two ways in determining the objectives of learning evaluation. The first

determines the details of the scope of the evaluation, the second, details the mental processes to be evaluated (Asrori et al., 2012). A lecturer or teacher must be able to pay attention and know the learning success achieved by his students. According to Djamarah, to find out the indicators of learning success can be seen from the absorption of students and the behavior that exists in students. Absorption is the level of mastery of the subject matter received by students individually or in groups (Supardi, 2015). Then the changes are in accordance with the basic competencies and indicators of learning success.

The indicator of learning outcomes if the learning outcomes achieved by students are in accordance with the criteria or values that have been set, using either the benchmark reference assessment or the norm reference assessment. Then, the teaching and learning process and learning achievement achieved by students are compared between before and after participating in teaching and learning activities or being given learning experiences.

As described in the previous chapter, that the purpose and urgency of developing a CEFR-based nahwu and shorof learning evaluation is in addition to knowing student learning outcomes with CEFR standards as well as diagnosing student difficulties, because CEFR in foreign language learning has become a standard in many countries (Nurdianto et al., 2020). Therefore, the hope of this research in the form of developing an evaluation of CEFR-based nahwu and shorof learning at MA An-Nur Jagasatru Cirebon can be used as a basic reference.

From the explanation above, it should be noted that in conducting the assessment, the teacher should pay attention to the assessment instrument with various instruments according to the characteristics of the competencies to be achieved by students. In assessing aspects of learning ability, it should be assessed comprehensively by considering various aspects.

It is possible to say that learning evaluation is an activity in the process of determining the level of achievement such as measurement, assessment, processing, interpretation, and consideration of student learning outcomes systematically. Learning evaluation is a process of gathering information that serves as the foundation for the level of development, planning, and achievement of the learning system in order to conform to the determined learning objectives.

CONCLUSIONS

Based on the research that has been done by the researchers, it can be concluded that the evaluation of CEFR-based nahwu and shorof learning for Ma An-Nur Jagasatru Cirebon students developed by the researchers has been declared valid by material experts and linguists. It is evident from the results of

the validation of material experts with a percentage of 85% from validators I and II, and linguists with an overall percentage of 88% from validators I and II, that they are suitable for using in nahwu and shorof learning. And based on the results of the analysis of the value of prototype 2, with an average value of 74.71, these results indicate that the CEFR-based evaluation of nahwu and shorof learning developed by researchers is feasible to be used in nahwu and shorof learning. This research is limited to the planning, development, and learning of nahwu and shorof required by the CEFR requirements for nahwu and shorof. Therefore, more in-depth research is needed to evaluate learning, especially in nahwu and shorof learning in order to obtain an evaluation of learning that is in accordance with the needs of students and the needs of the times.

ACKNOWLEDGMENT

BIBLIOGRAPHY

Ali, M. HF (2021). Transformasi dan Digitalisasi Pendidikan Dimasa Pandemi. In *Prosiding Seminar Nasional Pendidikan Program Pascasarjana Universitas PGRI Palembang* (pp. 15-16).

Ardiansyah, A. A., & Muhammad, A. (2020). Implementation of integrative Arabic grammar (Nahwu & Sharaf) curriculum in islamic boarding school. *Izdihar: Journal of Arabic Language Teaching, Linguistics, and Literature*, *3*(3), 211–238. https://doi.org/10.22219/jiz.v3i3.13264

Arikunto, S. (2021). *Dasar-dasar evaluasi pendidikan edisi III*. PT Bhumi Aksara.

Asrori, I., Thohir, M., & Ainin, M. (2012). *Evaluasi pembelajaran bahasa Arab*. Misykat.

Chaqoqo, S. G. N. (2017). Evaluasi pembelajaran nahwu dalam bentuk munaqasyah di PP Al-Luqmaniyyah Yogyakarta. *Lisania: Journal of Arabic Education and Literature*, *1*(1), 17–36. https://doi.org/https://doi.org/10.18326/lisania.v1i1.17-36

Dessiana Irma Yuanita, Hamdi Akhsan, K. W. (2015). Pengembangan panduan praktikum spektroskopi pada mata kuliah fisika modern. *Jurnal Inovasi dan Pembelajaran Fisika*, *2*(1), 77–87. https://doi.org/https://doi.org/10.36706/jipf.v2i1.2357

Egok, A. S., & Hajani, T. J. (2018). Pengembangan multimedia interaktif pada pembelajaran IPA bagi siswa sekolah dasar Kota Lubuklinggau. *Journal of Elementary School (JOES)*, *1*(2), 141–157. https://doi.org/10.31539/joes.v1i2.446

Foley, J. A. (2019). Adapting CEFR for English language education in ASEAN, Japan and China. *The New English Teacher*, *13*(2), 101–117.

He, L., & Chen, D. (2017). Developing common listening ability scales for Chinese learners of English. *Language Testing in Asia*, *7*(1), 1-12. https://doi.org/10.1186/s40468-017-0033-4

Hermanto, B. (2020). Perekayasaan sistem pendidikan nasional untuk mencerdaskan kehidupan bangsa. *FOUNDASIA*, *11*(2), 52-59. https://doi.org/10.21831/foundasia.v11i2.26933

Indayati, F., Saraswati, N. L., Rosalita, D. E., Imawan, J. S., & Dwi Febri. (2017). Ketimpangan Pendidikan dalam Perspektif Sosiologis di Indonesia. In *Ketimpangan dalam Pembangunan: Bunga Rampai Realitas Sosial atas Ketimpangan Pembangunan di Indonesia*. Media Nusa Creative.

Jamun, Y. M. (2018). The impact of technology on education. *Jurnal Pendidikan dan Kebudayaan Missio*, *10*(1), 48-52.

Jinani, H. (2013). *Pengembangan media evaluasi istima'dan qira'ah menggunaan hot potatoes 6 dalam pembelajaran bahasa Arab pada kelas VIII di MTs* (Skripsi, Universitas Negeri Malang).

Kadi, T., & Awwaliyah, R. (2017). Inovasi pendidikan: Upaya penyelesaian problematika pendidikan di Indonesia. *Jurnal Islam Nusantara*, *1*(2), 144-155. https://doi.org/10.33852/jurnalin.v1i2.32

Lubis, M. (2016). Peluang pemanfaatan pembelajaran berorientasi teknologi informasi di lingkup madrasah (mempersiapkan madrasah berwawasan global). *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah*, *1*(2), 147-153. https://doi.org/10.24042/tadris.v1i2.1063

Prasetya, M. A. (2021). Pengembangan evaluasi pembelajaran berbasis ispring pada mata pelajaran fisika pokok bahasan besaran dan satuan untuk sma kelas x. *Seminar Nasional Pendidikan IPA*, *Vol. 1 (1)*.

Maryana, N., & Liudmyla, B. (2019). Intercultural competence in CEFR: Implementation in practice. *The Scientific Issues of Ternopil Volodymyr Hnatiuk National Pedagogical University. Series: Pedagogy*, *0*(1). https://doi.org/10.25128/2415-3605.19.1.10

Ngafifi, M., & Ngafifi, M. (2014). Advances in Technology and Patterns of Human Life in Socio-Cultural Perspective. *Jurnal Pembangunan Pendidikan: Fondasi Dan Aplikasi*, *2*(1), 33-47. https://doi.org/10.21831/JPPFA.V2I1.2616

Nugraha, M. H. (2017). *Pengembangan Instrumen Pembelajaran Bahasa Arab berbasis Multimedia dengan Menggunakan Wondershare Quiz Creator*. UPI Press.

Nurdianto, T., Hidayat, Y., & Wulandari, V. A. (2020). CEFR-based Arabic language learning competency. *Izdihar: Journal of Arabic Language Teaching, Linguistics, and Literature*, *3*(3), 229-248. https://doi.org/10.22219/jiz.v3i3.14123

Nurkholis, N. (2013). Pendidikan dalam upaya memajukan teknologi. *Jurnal kependidikan*, *1*(1), 24-44. https://doi.org/10.24090/jk.v1i1.530

Permatasari, D. P. R. P., & Aliyah, D. F. A. (2021, December). Kualitas Buku Teks Pembelajaran Bahasa Arab Di MAN 2 Surakarta Kelas X Bermuatan Kurikulum 2013 (Studi Komparasi Antara Buku Teks Bahasa Arab Wajib dan Bahasa Arab Peminatan). In *Proceeding AEC (Arabic Education Conference)* (Vol. 1, No. 1, pp. 139-150).

Poupounaki-Lappa, P., Peristeri, T., & Coniam, D. (2021). Towards a communicative test of reading and language use for Classical Greek. *Journal of Classics Teaching*, *22*(44), 98-105. https://doi.org/10.1017/S2058631021000222

Basmar, E., Purba, B., Nugraha, N. A., Purba, E., Krisnawati, L., Damanik, D., ... & Siswanti, I. (2021). *Perekonomian dan bisnis indonesia*. Yayasan Kita Menulis.

Rahayu, E. E. (2014). Pengembangan alat evaluasi pembelajaran berbasis information and communication technologies (ICT) pada materi mengelola dokumen transaksi. *Jurnal Pendidikan Akuntansi (JPAK)*, *2*(2) 1-7. Retrieved from https://ejournal.unesa.ac.id/index.php/jpak/article/view/9139

Ridho, U. (2018). Evaluasi Dalam Pembelajaran Bahasa Arab. *An Nabighoh Jurnal Pendidikan dan Pembelajaran Bahasa Arab*, *20*(01), 19-26. https://doi.org/10.32332/an-nabighoh.v20i01.1124

Hasibuan, H. (2016). Studi kompetensi guru Pendidikan Agama Islam dalam pelaksanaan evaluasi pembelajaran. In *Forum Paedagogik, 8*(2), 14-38.

Sa'adah, N. (2019). Problematika pembelajaran nahwu bagi tingkat pemula menggunakan Arab pegon. *Lisanan Arabiya: Jurnal Pendidikan Bahasa Arab*, *3*(1), 15-32. https://doi.org/10.32699/liar.v3i01.995

Setiawan, D. (2018). Dampak perkembangan teknologi informasi dan komunikasi terhadap budaya. *JURNAL SIMBOLIKA: Research and Learning in Communication Study (E-Journal)*, *4*(1), 62-72. https://doi.org/10.31289/simbollika.v4i1.1474

Sibagariang, D., Sihotang, H., & Murniarti, E. (2021). Peran guru penggerak dalam pendidikan merdeka belajar di indonesia. *Jurnal Dinamika Pendidikan, 14*(2), 88-99.

Supardi. (2015). *Penilaian Autentik*. Raja Grafindo.

Nurdianto, T., & bin Ismail, N. A. (2020). Pembelajaran bahasa Arab Berbasis Common European Framework of Reference for Language (CEFR) di Indonesia. *Al Mahāra: Jurnal Pendidikan Bahasa Arab, 6*(1), 1-22. https://doi.org/10.14421/almahara.2020.061.01

Toriqul Arif, M. (2019). Penelitian evaluasi pendidikan. *Addabana : Jurnal Pendidikan Agama Islam, 2*(2), 66–76. https://doi.org/10.47732/adb.v2i2.122

Uri, N. F. M., & Aziz, M. S. A. (2020). The appropriacy and applicability of English assessment against CEFR global scale: Teachers' judgement. *3L: Language, Linguistics, Literature, 26*(3), 53-65. https://doi.org/10.17576/3L-2020-2603-05

Utomo, K., Soegeng, A. Y., Purnamasari, I., & Amaruddin, H. (2021). Pemecahan masalah kesulitan belajar siswa pada masa pandemi Covid19. *Mimbar PGSD Undiksha, 9*(1), 1-9. https://doi.org/10.23887/jjpgsd.v9i1.29923

Wahyono, I. (2019). Strategi kiai dalam mensukseskan pembelajaran nahwu dan shorof di Pondok Pesantren Al-Bidayah Tegalbesar Kaliwates Jember. *Jurnal Tarbiyatuna: Kajian Pendidikan Islam, 3*(2), 106-121. https://doi.org/10.29062/tarbiyatuna.v3i2.262

Westomi, J. A., Ibrahim, N., & Sukardjo, M. (2018). Pengembangan paket modul cetak mata pelajaran Pendidikan Agama Islam (PAI) untuk siswa SMA Negeri 1 Wangi-wangi Kabupaten Wakatobi. *JTP - Jurnal Teknologi Pendidikan, 20*(2), 138–151. https://doi.org/10.21009/jtp.v20i2.8628

Widiastuti, A., Rahelly, Y., & Sayfdaningsih, S. (2019). Pengembangan bahan ajar bentuk-bentuk geometri berbasis cerita untuk anak usia 5-6 tahun di TK Mazharul Iman Palembang. *Jurnal Pendidikan Anak, 7*(2), 176–189. https://doi.org/10.21831/jpa.v7i2.24461

Akhmadan, W. (2017). Pengembangan bahan ajar materi garis dan sudut menggunakan macromedia flash dan moodle kelas VII sekolah menengah pertama. *Jurnal Gantang, 2*(1), 27-40. https://doi.org/10.31629/jg.v2i1.62

Ng, H. Y., & Mohamad Zohir, A. (2021). Secondary school English teachers' knowledge and practice on CEFR-aligned English curriculum. *Asia Pacific Journal of Educators and Education, 36*(1), 75-91. https://doi.org/10.21315/apjee2021.36.1.5

Yoshitomi, A. (2021). Developing an assessment system of speaking skills in English as a foreign language based on CEFR-J. *Impact, 2021*(2), 56-58.

https://doi.org/10.21820/23987073.2021.2.56

Zazkia, S. A., & Hamami, T. (2021). Evaluasi kurikulum Pendidikan Agama Islam di tengah dinamika politik pendidikan di Indonesia. *AT-TA'DIB: Jurnal Ilmiah Prodi Pendidikan Agama Islam*, 82-93. https://doi.org/10.47498/tadib.v13i01.524

Zou, S., & Zhang, W. (2017). Exploring the adaptability of the CEFR in the construction of a writing ability scale for test for English majors. *Language testing in Asia*, *7*(1), 1-16. https://doi.org/10.1186/s40468-017-0050-3