



## MODEL OF ARABIC AUTHENTIC ASSESSMENT INSTRUMENTS: SPEAKING (*KALĀM*) AT MADRASAH ALIYAH

**Raswan**

Universitas Islam Negeri Syarif Hidayatullah Jakarta  
raswan@uinjkt.ac.id

**N. Lalah Alawiyah**

Universitas Islam Negeri Syarif Hidayatullah Jakarta  
lalah@uinjkt.ac.id

**Taufik Luthfi**

Sekolah Tinggi Agama Islam Dr. KH. EZ. Muttaqien  
taufikluthfi@staimuttaqien.ac.id

### **Abstract**

*Authentic assessment is a model for assessing learning processes and outcomes that differ from ordinary assessments. Authentic judgment must be done in a natural context, not contrived, applied and taught. Authentic assessment is also an assessment mandate in Indonesia's latest curriculum, namely the 2013 Revised Curriculum (revised K.13). The speaking aspect is at the core of successfully learning other aspects of Arabic language skills: listening (*istimā'*), reading (*qirā'ah*), writing (*kitābah*) and culture (*Ṣaqafah*). With proper speaking assessment will be the capital of success in learning speaking and other aspects of learning Arabic. Therefore, a, this article aims to present an authentic assessment instrument model of aspects of Arabic speaking learning (*kalām*) based on revision K.13. The research method used is research on the development of 4D models, define (definition), design (Design), develop (Development) and disseminate (trial). The results showed that the speech skills assessment instrument developed in this article has good validity, through expert tests (84%) and public tests (78%). The results of this study state that the developed speech ability assessment instrument is suitable for use.*

**Keywords:** *Assessment, authentic, kalām, valid.*

### **Abstrak**

Penilaian otentik adalah model untuk menilai proses dan hasil pembelajaran yang berbeda dari penilaian biasa. Penilaian otentik harus dilakukan dalam konteks alami, tidak dibuat-buat, diterapkan dan diajarkan. Penilaian otentik juga merupakan amanat penilaian dalam kurikulum terbaru Indonesia, yaitu Kurikulum Revisi 2013 (revisi K.13). Aspek berbicara adalah inti dari keberhasilan mempelajari aspek lain dari kemampuan bahasa Arab: mendnyimak (*istimā'*), membaca (*qirāah*), menulis (*kitābah*)

dan budaya (*Saqafah*). Dengan penilaian berbicara yang tepat akan menjadi modal keberhasilan dalam berbicara belajar dan aspek lain dari belajar bahasa Arab. Oleh karenanya, a, artikel ini bertujuan menyajikan model instrumen penilaian otentik aspek pembelajaran berbicara (kalam) bahasa Arab berdasarkan revisi K.13. Metode penelitian yang digunakan adalah penelitian pengembangan model 4D, *define (definition), design (Design), develop (Development) and disseminate (trial)*. Hasil penelitian menunjukkan bahwa instrumen penilaian keterampilan berbicara yang dikembangkan dalam artikel ini memiliki validitas yang baik, melalui uji ahli (84%) dan uji public (78%). Hasil penelitian ini menyatakan instrument penilaian kemampuan berbicara yang dikembangkan layak digunakan.

**Kata Kunci:** *Penilaian, otentik, kalām, valid.*

## INTRODUCTION

Learning Arabic is still considered difficult<sup>1</sup>. There are many problems: curriculum, teachers, media, infrastructure, strategies, even evaluation instruments. Arabic assessments occupy the most concerning position: UTS and UAS held by schools only test silent reading skills while other aspects are very untouched. Authentic assessment is the key, especially the aspect of speaking, so in this article we will examine the authentic instrument model of speaking Arabic for MA.

One of the characters of K.13 is a typical and well-known assessment standard with authentic assessment and non-authentic (traditional assessment) which is a harmonious pair of scientific learning approaches as process standards in K.13<sup>2</sup>. The scientific approach will not be meaningful when it is not paired with authentic judgments and vice versa. The regulation governing the

<sup>1</sup> Nur Habibah, “Lingkungan Artifisial Dalam Pembelajaran Bahasa Arab,” *Arabiyat : Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban* 3, no. 2 (2016): 173–196; Asep Muhammad et al., “Faktor Demotivasi Pembelajaran Bahasa Arab,” *Muhammad Asep Saeful Islam* 2, no. 1 (2015): 1–16; Rahmat Iswanto, “Pembelajaran Bahasa Arab Dengan Pemanfaatan Teknologi,” *Arabiyatuna : Jurnal Bahasa Arab* 1, no. 2 (2017): 139; Raswan Raswan, “Hisāb Ikhtibār Al-Lughah Al-‘Arabiyah ‘Alā Al-Mustawa Al Wathany Li L-Madāris as-Tsānawiyah Al-Islāmiyyah Fī Indonesiā,” *Jurnal Pendidikan Islam* 3, no. 1 (2017): 95; Raswan Raswan, “Tathwīr Namūdżaj Al-Muqarrar Al-Dirāsī Li Mādah Tashmīm Tadrīs Al-Lughah Al-‘Arabiyah ‘Alā Asās Al-Nasyāthāt,” *Arabi: Journal of Arabic Studies* 2, no. 1 (2017): 212–228; Siti Sulaikho and Lailatul Mathoriyah, “Analisis Permasalahan Mahasiswa Dalam Mempelajari Sintaksis Bahasa Arab,” *Jurnal Education and Development* 8, no. 3 (2020): 293–297; Shafruddin Tajuddin, “Pengembangan Model Pembelajaran Bahasa Arab Tingkat Sekolah Dasar Untuk Meningkatkan Kemampuan Berbahasa Arab Siswa,” *PARAMETER: Jurnal Pendidikan Universitas Negeri Jakarta* 29, no. 2 (2017): 200–215.

<sup>2</sup> Deprizon, “Pengembangan Pembelajaran Hifzhil-Qur’an Dengan Penilaian Autentik Sebagai Penggerak Karakteristik Siswa Di Lembaga Pendidikan Formal,” *Kreatifitas Jurnal Ilmiah Pendidikan Islam* 10, no. 1 (2021): 22–35, <https://ojs.diniyah.ac.id/index.php/Kreatifitas/article/view/270/190>; Latifatul Muyasaroh, “Implementasi Penilaian Autentik Kelas 1 Pada Pembelajaran Tematik, Tema: 4 Keluargaku Di SD Negeri 01 Purwanegara Kec. Purwokerto Utara Kab. Banyumas” (IAIN Purwokerto, 2019); Naeli Sangadah, “Penilaian Autentik Pada Mata Pelajaran SKI Di MI Darul Hikmah Bantarsoka Kec. Purwokerto Barat Kab. Banyumas” (IAIN Purwokerto, 2016); Zaenatun Nafisah, “Implementasi Penilaian Autentik Pada Mata Pelajaran Ppkn Kelas V Di Madrasah Ibtidaiyah Negeri Purwokerto Tahun Pelajaran 2016/2017” (IAIN Purwokerto, 2017); Raswan Raswan, “Dialek Laki-Laki Dan Perempuan Dalam Konteks Bahasa Arab,” *alMahara Jurnal Pendidikan Bahasa Arab* 5, no. 2 (2019): 271–300; Harry Soeprianto Muhammad Turmuzi, Nyoman Sridana, Ketut Sarjana, “Peningkatan Kemampuan Guru Sekolah Dasar Di Gugus II Kecamatan Lembar Dalam Menerapkan Authentic Assesment Dalam Pembelajaran,” *Jurnal Pendidikan dan Pengabdian Masyarakat* 2, no. 4 (2019): 472–479.

obligation to use this authentic assessment is permendikbud no. 66 of 2013 which was later refined with the birth of Permendikbud no. 104 of 2014, while the standard setting process is set out in Permendikbud No. 65 of 2013.

Authentic assessment requires a complete understanding of the teacher as the spearhead of the curriculum. Without a comprehensive understanding of the teacher, this concept will be meaningless. Even more worrying is the fact that shows that a weak standard of education among madrasas and Arabic language teachers in particular is the field of educational assessment. This was seen in the author's visitation in 2013 when researchers conducted an assessment of madrasas in the South Jakarta area against accreditation documents before being visited / assessed by a team from the School / Madrasah Accreditation Agency or known as BAN-S / M.

In developing the Arabic syllabus there are significant differences between Madrasah Ibtidaiyah (MI) and Madrasah Aliyah (MA) with Madrasah Tsanawiyah (MTs)<sup>3</sup>. In the MTs syllabus, emphasis on cognitive aspects and skills is presented separately and based on temporary skills in MI and MA knowledge and skills are merged in learning. If MTs Syllabus is based on proficiency, then in MI and MA based on a combination of two or more Arabic skills. Therefore an authentic assessment instrument developed in MTs is based on one skill while in MI and MA two or more skills in one instrument. But in this article with all the limitations it will only focus on speaking skills in MA.

Authentic assessment is seen as the oldest assessment in education and learning according to the Arabic version. Because since the time of ignorance this authentic assessment has been used for example in assessing the quality of the poetry of poets, writers, in displays posted in markets such as the Ukadz market, as well as in riding and exercising. As a result of the literary work, the seven best posted on the Ka'bah are known as "المعلقات السبع". While in its modern study authentic assessment is considered to have begun to appear since the eighth decade of the twentieth century,

---

<sup>3</sup> Azkia Muharom Albantani, "Implementasi Kurikulum 2013 Pada Pembelajaran Bahasa Arab Di Madrasah Ibtidaiyah," *Arabiyat: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 2, no. 2 (2015): 178–191; Muhammad Mas'ud, "Konsep Life Skills Dalam Pembelajaran Bahasa Arab Di Madrasah," *LISANIA: Journal of Arabic Education and Literature* 1, no. 2 (2017): 20–39; Abdul Mutholib, "ANALISIS BUTIR SOAL BAHASA ARAB UJIAN AKHIR MADRASAH BERSTANDAR MADRASAH IBTIDAIYAH ( MI )," *Arabia: Jurnal Pendidikan Bahasa Arab* 5, no. 2 (2013): 141–161; Raswan, "Tathwîr Namûdzaj Al-Muqarrar Al-Dirâsi Li Mâdah Tashmîm Tadrîs Al-Lughah Al-'Arabiyyah 'Alâ Asâs Al-Nasyâthât"; Agus Tricahyo, "Landasan Filosofis Kebijakan Pengembangan Kurikulum Bahasa Arab," *Cendekia: Jurnal Kependidikan Dan Kemasyarakatan* 11, no. 1 (2013): 57–74; Ahmad Zubaidi, "Model-Model Pengembangan Kurikulum Dan Silabus Pembelajaran Bahasa Arab," *Cendekia: Jurnal Kependidikan Dan Kemasyarakatan* 13, no. 1 (2015): 107–122.

since it was first applied in the United States, with the name The Assessment of Performance Unit (APU)<sup>4</sup>

Authentic assessment in the modern world was born because of a change in the concept of education from behaviorism towards the theory of constructivism / Constructivist Learning Theory (نظرية "البنوية" لعملية التعلم), then a new assessment method and approach was known as Kinejra-Based Assessment / Performance Based Assessment (التقييم في ضوء الأداء) or Authentic Assessment / Authentic Assessment (التقييم الأصيل / الحقيقي / الواقعي), the more popular term for this latest assessment in Arabic is التقييم البديل which is literally interpreted as an alternative assessment<sup>5</sup>

Meanwhile, as quoted by Palm from Kirst, explained about the use of the term authentic assessment, namely “use the term of authentic assessment is intended to convey that the assessment tasks themselves are real instances of extended criterion performances, rather than proxies or estimators of actual learning goals. other synonym are direct or performance assessment” (Palm, tt: 7) that the use of the term authentic appraisal is intended to convey that the task in the assessment itself is a concrete example that is elaborated by performance criteria, rather than a proxy or estimator / estimate of the actual learning objectives. So the synonym of authentic assessment is actually a direct assessment or assessment of real performance examples.

In Arabic authentic assessment is defined as "التقييم غير المقيد بنتائج امتحانات تقريرية أو اختبارات مقننة." which is a judgment that is not determined by the standardized test and is an in-depth, regular, continuous and daily assessment<sup>6</sup>.

Formally, in Permendikbud no. 66 of 2013<sup>7</sup> (Standar Penilaian Pendidikan, 2013) explained that Educational Assessment as a process of gathering and processing information to measure the achievement of student learning outcomes includes: authentic assessment, self-assessment, portfolio-based assessment, repeat, daily repetition, midterm replication, end semester test, competency level test, competency level quality test, national exam, and school / madrasah exam. It was further explained that authentic assessment is a comprehensive assessment to assess starting from input, process, and output of authentic learning and assessment carried out by the teacher on

<sup>4</sup> A Al-Syahri wa A 'Ubadah, *Al-Tawjihat Al-Haditsah Li Taqvim Al-Thullab Li Maddat Qira'at Fi Al-Takhashshush/Dukturah/Al-Mustawa Al-Tsalits* (Wizarat al Ta'lim al-'Ali wa Jami'ah al-Imam Muhammad bin Sa'ud al-Islami, n.d.), 50.

<sup>5</sup> S. A. Al-Mujahid, “Nahw Ru'yah Jadidah Li Ishlah Midzam Al-Qiyas Wa Al- Taqvim Al-Ta'limi Fi Libiya,” *Al-Majallah al-Jami'ah* (n.d.): 15.

<sup>6</sup> Ahmad Abu Zainuddin, “Al-Taqvim Al-Badil Fi Al-Lughah Al-'Arabiyyah Li Talamidz Al- Marhalah Al-Ibtidaiyyah,” last modified 2004, Bouzein@aub.Edu.Lb.

<sup>7</sup> Tim Penyusun, *Peraturan Menteri Pendidikan Dan Kebudayaan Tentang Standar Penilaian Pendidikan*, 2013.

an ongoing basis. So authentic assessment is one type of educational assessment, has a major position compared to other types of assessment and can only be done by a teacher, meaning that if there is no teacher, authentic assessment cannot be done.

So authentic assessment is a comprehensive assessment (input-process-output, attitude-knowledge-skill), meaningful, reflecting the real world (the assignment is a miniature of the real world in everyday life), using a variety of strategies and assessment instruments and has different characteristics from the usual assessment.

The development of authentic assessment instruments is important to be a guide for teachers to be able to develop valid assessments. Several studies have been conducted on the development of authentic assessments of the Arabic language<sup>8</sup> calm Authentic Assessment in Arabic Language Learning,<sup>9</sup> on Authentic Learning Evaluation and Assessment in Arabic Language Learning, and<sup>10</sup> about the Development of Curriculum-Based Authentic Assessment Instruments 2013 in the Arabic Language Textbook Class X Semester 1 Madrasah Aliyah. These studies have not examined the development of authentic assessment instruments for Arabic subjects in terms of speech ability (al-kalam).

Therefore, this research was conducted to develop an authentic assessment model in the assessment of speech ability in Arabic language learning. The results of this study will be useful for practitioners in developing authentic assessment instruments in carrying out more optimal Arabic language learners.

## **METHOD**

This research uses the research steps of 4D (four-D Model) development. Activities in this model consist of define (definition), design (Design), develop (Development) and disseminate (trial). At the Define (defining) stage researchers conducted theoretical studies, interviews with experts, the best teachers in Indonesia specifically in the field of Arabic, related ministries of religion, related ministries, documentation studies, observations in the field and others. The selected teachers are Arabic teachers who are often involved in the development of curriculum K.13 and the creation of teaching books and student books K.13. The experts in question are lecturers and/or professors who concentrate in the learning of Arabic and Akhlak Creed.

---

<sup>8</sup> Salamah; Rifqi Aulia Rahman; M. Elfan Kaukab, "Penilaian Autentik Dalam Pembelajaran Bahasa Arab," *NIVEDANA: Jurnal Komunikasi Dan Bahasa* 2, no. 1 (2021): 60–75.

<sup>9</sup> Dina Indriana, "Evaluasi Pembelajaran Dan Penilaian Autentik Dalam Pembelajaran Bahasa Arab," *Al-Ittihad: Jurnal Keilmuan Dan Kependidikan Bahasa Arab* 10, no. 2 (2018): 34–52.

<sup>10</sup> Resi Agustien, "Pengembangan Instrumen Penilaian Autentik Berbasis Kurikulum 2013 Pada Buku Ajar Bahasa Arab Kelas X Semester 1 Madrasah Aliyah," *La-Tahzan: Jurnal Pendidikan Islam* 12, no. 1 (2020): 52–70.

The design phase analyzed KI/KD Arabic class X while mapping it into KD groups complete with learning indicators, then selected only KD on KI 1, 2 and 3 as the object of the study. After that researchers designed observation instruments and performance assessments of KD and these indicators. Furthermore, the development stage, researchers make revisions and re-corrections to the instruments that have been made, showing the instruments to experts both Arabic to be given input. From this input, the researcher made revisions. After revision the instrument is duplicated on a limited basis.

The last stage, disseminate (dissemination) in the form of trials and publications, researchers spread the questionnaire to assess the instrument to limited respondents namely four experts and five practitioner teachers in Jakarta and Banten. For reliable results, researchers conducted field observations looking at Arabic learning in class X both in MAN and in MAS. After the questionnaire is filled the author performs analysis and revision of the instrument to then be replaced more for wider public testing. In this limited and broader public test, researchers conducted two analyses: qualitative and quantitative analysis. Qualitative analysis is useful for instrument improvement while quantitative analysis is used for experimental tests between experimental instruments and control instruments i.e. those in teacher books and K.13 student books. Then the publication of both oral and written, oral publications are delivered in the presence of Arabic teachers so that Arabic teachers get exposure to authentic studies specifically for arabic in-depth more.

## RESULT AND DISCUSSION

Based on the results of literature studies, interviews with experts, the best teachers in Indonesia specialized in the field of Arabic language involved in the development of the national curriculum, related religious ministries, observations in the field, documentary studies of several Aliyah Madrasas which among the teachers were madrasa curriculum development teams specialized in Arabic language such as MAN 1 Suarakarta, MAN 1 Bojonegoro, MAN 1 Rengel, Tuban, MAN 3 Malang, MAN 1 Malang, MA Pembangunan UIN Jakarta and MAN 16 Jakarta, researchers designed assessment instruments especially skills. The initial product that was designed was an instrument of assessment of Arabic language skills Madrasah Aliyah. The following are examples of instruments designed by researchers:

### Initial Instrument Design

The following are examples of instruments designed by researchers:

Table 1. KD and Arabic Indicators (KI-4)

KD	Indikator
4.1 Conduct a simple dialogue according to the context appropriately and smoothly regarding the topic: البيانات الشخصية؛ المرافق العامة في المدرسة؛ الحياة by paying attention to linguistic elements, the structure of the text and cultural elements correctly and in context	4.1.1. Recite the sounds of letters, words, phrases and sentences 4.1.2. Demonstrate various oral and written information
4.2 Delivering various simple oral information about: البيانات الشخصية؛ المرافق العامة في المدرسة؛ الحياة by paying attention to linguistic elements, the structure of the text and cultural elements correctly and in context	

**Skills Assessment Instrument:**

**Instructions:**

Have a dialogue with some of your friends regarding the topics discussed.

**Reporting Guidelines:**

Nama Siswa	Tes Dialog (1-5)										Jml (10-50)
	ID	T	PD	G	F	L	V	M	I	S	

Ket: ID: Isi Dialog, T: Tampilan, PD: Percaya Diri, G: Gestur, F: Fasih, V: Volume, M: Makhraj, I : Intonasi, S: Stressing.

**Early Design Chord Gudelines Arabic Skills 1**

**Instructions:**

Show in front of the class to explain the contents related to the topics discussed using Arabic.

**Suspension Guidelines:**

Nama Siswa	Tes Pidato (1-5)										Jml (10-50)
	IP	T	PD	G	F	L	V	M	I	S	

Ket: IP: Isi Pidato, T: Tampilan, PD: Percaya Diri, G: Gestur, F: Fasih, L: Lancar, V: Volume, M: Makhraj, I: Intonasi, S: Stressing.

## **Preliminary Design Guidelines for The Suspension of Arabic Skills 2 First Test (Expert Test)**

The first test in question is a test of Arabic education experts. Examiners of Arabic language experts consist of four lecturers with a final education in S2 and are in the process of completing S3 and one of the doctorates from India. The three doctoral candidates in question are AR (doctoral candidates for PBA UIN Malang), UR and WS (both doctoral candidates concentrating on PBA S.Ps UIN Jakarta while the doctorate in question is AD. These four experts in the field of PBA, education and Arabic content are considered qualified.

All four of them provide qualitative and quantitative values, quantitative values are useful for producing numerical data regarding the quality of designs that have been developed by researchers. While the qualitative assessment provides real input on the proposed improvement of instruments that have been developed by researchers.

Among the proposals is to make the trip four so that it is easier to apply it in Madrasah by the teacher, besides that, some language and term errors do not fail to be part of entering these experts besides the choice of language in the indicators of learning achievement cited from the syllabus is too general especially in indicator 4.2.1. Demonstrate various oral and written information. In more detail, AR provides input that the scale should be only 4 in order to facilitate the teacher in carrying out the assessment, intonation is the same as the volume, the same as fluency, and dialogue and speech tests are actually linguistic tests.

While quantitative assessment of the instruments designed by the researcher, quantitative data is generated; first, Arabic assessment instruments get the average value of each examiner; WS gives a value with a mean of 8, AR gives a mean value of 11, AD gives a mean value of 10, and UR gives a mean value of 8, so the average value of each instrument is 9.8. The following is an overview of the results of the assessment:

Table 2. Arabic Expert Test Score

<b>Expert Name</b>	<b>Value of Arabic Skill Instrument</b>
AR	11
WS	8
AD	10
UR	8

Jumlah	37
Rata-rata	84%

Based on the input from experts, the researchers revised the first stage of the product into the following:

### Arabic Expert Test Revision Table 1

The resulting instruments are as follows:

Assessment Strategy : Performance Test

Instruments : Observation Guidelines

Scale : Thurstone

Assessment Aspect : Reciting

**Instructions:** Recite the sounds of letters, words, phrases and sentences with some of your friends regarding the topics discussed.

#### Suspension Guidelines:

Name Student	Dialog Test (1-4)										Total (10-40)
	ID	T	PD	G	F	L	V	M	I	S	

Ket: ID: Isi Dialog, T: Tampilan, PD: Percaya Diri, G: Gestur, F: Fasih, V: Volume, M: Makhraj, I: Intonasi, S: Stressing.

### Arabic Expert Test Revision Table 2

**Instructions:** Show in front of the class to demonstrate verbal and written information related to the topics discussed using Arabic.

#### Chord Gudelines:

Nama Siswa	Tes Pidato (1-4)										Jml (10-40)
	IP	T	PD	G	F	L	V	M	I	S	

Ket: IP: Isi Pidato, T: Tampilan, PD: Percaya Diri, G: Gestur, F: Fasih, L: Lancar, V: Volume, M: Makhraj, I: Intonasi, S: Stressing.

### Limited Public Test

Quantitative testing of instruments designed by researchers and assessed by experts or experts, quantitative data is produced; first, the Arabic assessment instrument obtaining the average value of each examiner for Arabic skill instruments is that ZA gives a mean value of 8.71 and 8.45,

YF gives a value with a mean of 6.64 and 62, FS gives a value with mean 8.36 and 7.35, and E gave mean values of 9.79 and 10.8 respectively and F with an average value of 9.21 and 8.3. The following is an overview of the assessment results:

Table 3. Arabic Limited Public Test Score (Experiment)

Expert Name	Arabic Skills Instrument Value
ZA	8.71
YF	6.64
FS	8.36
E	9.79
F	9.21
Rata-rata	8.54
Prosentase	78%

The value of Arabic language assessment skills each gets an average of 78%, respectively. While an assessment of the Arabic language assessment skills in the teacher's book and student's book gets the scores as in the following description:

Table 4. Arabic Limited Public Test Score (Controls)

Teacher's Name	Arabic Skills Instrument Score
ZA	8.21
YF	6
FS	7.71
E	7.93
F	7
Rata-rata	7.37
Prosesntase	67%

\* Note: from One of MAN Jakarta or MA at Pamulang, South Tangerang.

The value of Arabic assessment instruments is 67%, respectively. If calculated by the t test as the formula is in Chapter III, the significance for the Arabic language skills developed is 3.27 while  $t_{tabel}(\alpha 0,05) = 2.35$  then  $H_0$  rejected thus the instrument developed by the researcher is better than the control instrument. Likewise on non-Arabic test instruments (observation guidelines) get  $t_{hitung} = 3.061$  while  $t_{tabel}(\alpha 0,05) = 2.35$  then  $H_0$  rejected thus the observation guideline instrument developed by the researcher is better than the control instrument.

### Product Revision (مراجعة المنتج/الإنتاج)

From various inputs in Phase II testing, the researchers revised the product assessment skills to be as follows.

### Arabic Public Test Revision (Performance 1)

Skills

Assessment Strategy : Performance Test

Instruments : Observation Guidelines

Scale : Thurstone

Assessment Aspect : Conducting dialogue in Arabic

**Instructions** : Have a dialogue with some of your friends regarding the topics discussed.

#### Chord Gudelines:

Nama Siswa	Tes Dialog (1-4)										Jml (10-40)
	ID	T	PD	G	F	L	V	M	I	S	
Rahman											
Hakim											

Ket: ID: Isi Dialog, T: Tampilan, PD: Percaya Diri, G: Gestur, F: Fasih, V: Volume, M: Makhraj, I : Intonasi, S: Stressing.

### Arabic Public Test Revision (Performance 2)

**Instruksi:** Show in front of the class to explain the contents related to the topics discussed using Arabic.

#### Chord Gudelines:

Nama Siswa	Tes Pidato (1-4)										Jml (10-40)
	IP	T	PD	G	F	L	V	M	I	S	

Ket: IP: Isi Pidato, T: Tampilan, PD: Percaya Diri, G: Gestur, F: Fasih, L: Lancar, V: Volume, M: Makhraj, I: Intonasi, S: Stressing.

### Product Improvement (تكملة المنتج)

The following is a refinement of products that have been developed and exceeded the field testing process:

#### Arabic Skills Instrument

#### Arabic Public Test Final Product (Performance 1)

Assessment Strategy : Arabic Performance Test

Instruments : Observation Guidelines

Scale : Likert

Assessment Aspect : Pronounce the sounds of letters, words, phrases and sentences

**Instructions:** Recite the sounds of letters, words, phrases and sentences with some of your friends regarding the topics discussed.

### Chord Gudelines:

Nama Siswa	Tes Dialog (1-4)										Jml (10-40)
	ID	T	PD	G	F	L	V	M	I	S	
Rahman											
Hakim											

Ket: ID: Isi Dialog, T: Tampilan, PD: Percaya Diri, G: Gestur, F: Fasih, V: Volume, M: Makhraj, I : Intonasi, S: Stressing.

- The teacher can focus on certain aspects that are considered as core aspects.

### Arabic Public Test Final Product (Performance 2)

**Instruksi:** Show in front of the class to demonstrate verbal and written information related to the topics discussed using Arabic.

### Chord Gudelines:

Nama Siswa	Tes Pidato (1-4)										Jml (10-40)
	IP	T	PD	G	F	L	V	M	I	S	

Ket: IP: Isi Pidato, T: Tampilan, PD: Percaya Diri, G: Gestur, F: Fasih, L: Lancar, V: Volume, M: Makhraj, I : Intonasi, S: Stressing.

- The teacher can focus on certain aspects that are considered as core aspects.

### Arabic End Product

The product developed in this study consists of skills assessment instruments. Products that have been developed beyond the four main phases include the defining, designing, developing and finally disseminating phases. The product that has been produced by the researcher is only a model, where the teacher can reduce it, adding even replacing it if it is seen that this model is inappropriate with the conditions in the field. The final results of this research product are separated between each competency, but in the real field as in RPP and the implementation of learning for each competency is a unity that cannot be separated even coupled with assessment of other aspects that are not discussed in this study.

### CONCLUSION

Authentic speaking assessment instruments must be developed well and seriously so that assessment is not only a tool for assessing but also part of an Arabic learning strategy. The assessment instruments developed are adjusted to the situation and conditions of students, environment and infrastructure and infrastructure in schools / madrasas. The instrument must also be in accordance with the official curriculum issued by the ministry of religion. So that the instrument is in accordance with the existing KI / KD. Besides

having to go beyond the steps of instrument development from the start of making grids to reviewing and even testing and analyzing questions in the field if possible. The instrument of assessment of listening skills developed in this article has good validity both through expert tests and through field tests. So that schools / madrasas can make it a model and guide in developing istima assessment instruments. The results showed that the speech skills assessment instrument developed in this article has good validity, through expert tests (84%) and public tests (78%). The results of this study state that the developed speech ability assessment instrument is suitable for use. However, this model needs to be continuously updated, adjusted to the development, condition and conditions and facilities and infrastructure.

## REFERENCES

- Abu Zainuddin, Ahmad. "Al-Taqwim Al-Badil Fi Al-Lughah Al-'Arabiyyah Li Talamidz Al-Marhalah Al-Ibtidaiyyah." Last modified 2004. Bouzein@aub.Edu.Lb.
- Agustien, Resi. "Pengembangan Instrumen Penilaian Autentik Berbasis Kurikulum 2013 Pada Buku Ajar Bahasa Arab Kelas X Semester 1 Madrasah Aliyah." *La-Tahzan: Jurnal Pendidikan Islam* 12, no. 1 (2020): 52–70.
- Al-Mujahid, S. A. "Nahw Ru'yah Jadidah Li Ishlah Midzam Al-Qiyas Wa Al- Taqwim Al-Ta'limi Fi Libiya." *Al-Majallah al-Jami'ah* (n.d.): 15.
- Al-Syahri, A, A 'Ubadah. *Al-Tawjihah Al-Haditsah Li Taqwim Al-Thullab Li Maddat Qira'at Fi Al-Takhashshush/Dukturah/Al-Mustawa Al-Tsalits*. Wizarat al Ta'lim al-'Ali wa Jami'ah al-Imam Muhammad bin Sa'ud al-Islami, n.d.
- Albantani, Azkia Muharom. "Implementasi Kurikulum 2013 Pada Pembelajaran Bahasa Arab Di Madrasah Ibtidaiyyah." *Arabiyat: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 2, no. 2 (2015): 178–191.
- Deprizon. "Pengembangan Pembelajaran Hifzhil-Qur'an Dengan Penilaian Autentik Sebagai Penggerak Karakteristik Siswa Di Lembaga Pendidikan Formal." *Kreatifitas Jurnal Ilmiah Pendidikan Islam* 10, no. 1 (2021): 22–35. <https://ojs.diniyah.ac.id/index.php/Kreatifitas/article/view/270/190>.
- Indriana, Dina. "Evaluasi Pembelajaran Dan Penilaian Autentik Dalam Pembelajaran Bahasa Arab." *Al-Ittijah: Jurnal Keilmuan Dan Kependidikan Bahasa Arab* 10, no. 2 (2018): 34–52.
- Iswanto, Rahmat. "Pembelajaran Bahasa Arab Dengan Pemanfaatan Teknologi." *Arabiyatuna : Jurnal Bahasa Arab* 1, no. 2 (2017): 139.
- Latifatul Muyasaroh. "Implementasi Penilaian Autentik Kelas 1 Pada Pembelajaran Tematik, Tema: 4 Keluargaku Di SD Negeri 01 Purwanegara Kec. Purwokerto Utara Kab. Banyumas." IAIN Purwokerto, 2019.
- Mas'ud, Muhammad. "Konsep Life Skills Dalam Pembelajaran Bahasa Arab Di Madrasah." *LISANIA: Journal of Arabic Education and Literature* 1, no. 2 (2017): 20–39.
- Muhammad, Asep, Saepul Islam, Madrasah Aliyah, and Negeri Tanggeung. "Faktor Demotivasi Pembelajaran Bahasa Arab." *Muhammad Asep Saeful Islam* 2, no. 1 (2015): 1–16.

- Muhammad Turmuzi, Nyoman Sridana, Ketut Sarjana, Harry Soeprianto. “Peningkatan Kemampuan Guru Sekolah Dasar Di Gugus II Kecamatan Lembar Dalam Menerapkan Authentic Assesment Dalam Pembelajaran.” *Jurnal Pendidikan dan Pengabdian Masyarakat* 2, no. 4 (2019): 472–479.
- Mutholib, Abdul. “ANALISIS BUTIR SOAL BAHASA ARAB UJIAN AKHIR MADRASAH BERSTANDAR MADRASAH IBTIDAIYAH ( MI ).” *Arabia: Jurnal Pendidikan Bahasa Arab* 5, no. 2 (2013): 141–161.
- Nafisah, Zaenatun. “Implementasi Penilaian Autentik Pada Mata Pelajaran Ppkn Kelas V Di Madrasah Ibtidaiyah Negeri Purwokerto Tahun Pelajaran 2016/2017.” IAIN Purwokerto, 2017.
- Nur Habibah. “Lingkungan Artifisial Dalam Pembelajaran Bahasa Arab.” *Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban* 3, no. 2 (2016): 173–196.
- Raswan, Raswan. “Dialek Laki-Laki Dan Perempuan Dalam Konteks Bahasa Arab.” *alMahara Jurnal Pendidikan Bahasa Arab* 5, no. 2 (2019): 271–300.
- . “Hisāb Ikhtibār Al-Lughah Al-‘Arabiyyah ‘Alā Al-Mustawa Al Wathany Li L-Madāris as-Tsānawiyah Al-Islāmiyyah Fī Indonesiā.” *Jurnal Pendidikan Islam* 3, no. 1 (2017): 95.
- . “Tathwīr Namūdzaj Al-Muqarrar Al-Dirāsī Li Mādah Tashmīm Tadrīs Al-Lughah Al-‘Arabiyyah ‘Alā Asās Al-Nasyāthāt.” *Arabi: Journal of Arabic Studies* 2, no. 1 (2017): 212–228.
- Salamah; Rifqi Aulia Rahman; M. Elfan Kaukab. “Penilaian Autentik Dalam Pembelajaran Bahasa Arab.” *NIVEDANA: Jurnal Komunikasi Dan Bahasa* 2, no. 1 (2021): 60–75.
- Sangadah, Naeli. “Penilaian Autentik Pada Mata Pelajaran SKI Di MI Darul Hikmah Bantarsoka Kec. Purwokerto Barat Kab. Banyumas.” IAIN Purwokerto, 2016.
- Sulaikho, Siti, and Lailatul Mathoriyah. “Analisis Permasalahan Mahasiswa Dalam Mempelajari Sintaksis Bahasa Arab.” *Jurnal Education and Development* 8, no. 3 (2020): 293–297.
- Tajuddin, Shafruddin. “Pengembangan Model Pembelajaran Bahasa Arab Tingkat Sekolah Dasar Untuk Meningkatkan Kemampuan Berbahasa Arab Siswa.” *PARAMETER: Jurnal Pendidikan Universitas Negeri Jakarta* 29, no. 2 (2017): 200–215.
- Tim Penyusun. *Peraturan Menteri Pendidikan Dan Kebudayaan Tentang Standar Penilaian Pendidikan*, 2013.
- Tricahyo, Agus. “Landasan Filosofis Kebijakan Pengembangan Kurikulum Bahasa Arab.” *Cendekia: Jurnal Kependidikan Dan Kemasyarakatan* 11, no. 1 (2013): 57–74.
- Zubaidi, Ahmad. “Model-Model Pengembangan Kurikulum Dan Silabus Pembelajaran Bahasa Arab.” *Cendekia: Jurnal Kependidikan Dan Kemasyarakatan* 13, no. 1 (2015): 107–122.