

## **ANALISIS KUALITAS TES BAHASA ARAB BERBASIS *HIGHER ORDER THINKING SKILL (HOTS)***

**Indah Rahmi Nur Fauziah**

*FPBS, Universitas Pendidikan Indonesia*

*indahrahmi@student.upi.edu*

**Syihabudin**

*FPBS, Universitas Pendidikan Indonesia*

*syihabuddin@upi.edu*

**Asep Sopian**

*FPBS, Universitas Pendidikan Indonesia*

*asepsopian@upi.edu*

### **Abstract**

*Educational evaluation is an important aspect, because with educational evaluation, it can be known to what extent the educational process has been achieved, it can also determine the next step regarding what is to be done. In the realm of evaluation, analyzing the quality of HOTS-based tests is still rarely to be done as was the case in Lembang MTs Al-Musyawah. Arabic Language Teachers have not analyzed the quality of Arabic question items in the final examination (UAS). This research aims to determine the quality of hots-based instruments in final examination (UAS) questions of Arabic that are have not yet know. The type of research that used in this research is descriptive research, and 30 people as the sample of data. The results of this research show that the validity of the test be include high, and high reliability, 25 questions are not fit with the criteria for multiple choice writing, the level of difficulty not yet proporsional, distinguishing power enough, and effectiveness of deceivers are enough categories, and all questions include as Lots category with C1 numbers as much as 2 items, 2 items for C2, and C3 as much as 36 items of questions.*

*Keywords : evaluation, hots, test*

### **Abstrak**

*Evaluasi pendidikan merupakan aspek yang penting, karena dengan evaluasi pendidikan dapat diketahui sejauh mana proses pendidikan telah tercapai, juga dapat menentukan langkah selanjutnya mengenai apa yang hendak dilakukan. Dalam ranah evaluasi, melakukan penganalisisan terhadap kualitas tes berbasis HOTS masih sangat minim dilakukan seperti yang terjadi di MTs Al-Musyawah Lembang. Guru Bahasa Arab belum melakukan penganalisisan terhadap kualitas butir soal UAS Bahasa Arab. Penelitian ini bertujuan untuk mengetahui kualitas tes berbasis hots dalam soal UAS bahasa arab yang belum diketahui. Jenis penelitian yang digunakan adalah penelitian deskriptif. dengan jumlah sampel 30 orang. Hasil dari penelitian ini menunjukkan bahwa validitas tes termasuk tinggi, reliabilitas tinggi, 25 soal tidak sesuai kriteria kaidah penulisan pilihan ganda, tingkat kesukaran belum proporsional, daya pembeda sedang, dan efektifitas pengecoh termasuk kategori cukup, dan seluruh soal termasuk kategori Lots dengan angka C1 sebanyak 3 butir soal, C2 1 butir soal dan C3 sebanyak 36 butir soal.*

**Kata Kunci :** *Evaluasi, Hots, Tes*

### مستخلص البحث

يعد التقييم جانباً مهماً لأنه يمكن به أن تعرف جودة العملية التعليمية التامة. كما يمكنه تحديد الخطوات التالية فيما يتعلق بما سيفعل في التعليم. تحليل جودة الاختبار على أساس قدرة التفكير العليا نادر فعل في عملية التقييم، كما في مدرسة الثانوية المشاورة ليمبانج لم يحلل معلم اللغة العربية لجودة بنود الأسئلة لإمتحان الفصل النهائي باللغة العربية. يهدف هذا البحث لمعرفة جودة اختبار على أساس قدرة التفكير العليا والسفلى في أسئلة إمتحان الفصل النهائي باللغة العربية التي لم تُعرف. البحث المستعمل هو البحث الوصفي مع عينة من ٣٠ التلاميذ. تشير نتائج هذه البحث إلى أن صدق الاختبار تشمل عالية وثبات عالية ، وهناك ٢٥ سؤالاً لا تتوافق لمعايير قواعد الكتابة متعددة الخيارات، ومستوى الصعوبة ليست متناسبة و قوة التفرقة تشمل معتدلاً، وفعالية الخداع تشمل فئات كافية ، وتشمل جميع الأسئلة فئات قدرة التفكير السفلى مع C١ ٣ السؤال، C2 ١ السؤال و C3 ٣٦ سؤالاً.

الكلمات الرئيسية : تقييم, قدرة التفكير العليا, اختبار

## PENDAHULUAN

### A. Latar Belakang Masalah

Pada kegiatan evaluasi, dibutuhkan instrument untuk mengevaluasi suatu proses pembelajaran. Instrument tersebut dapat berupa tes maupun non tes. Teknik non tes diantaranya skala bertingkat, kuisisioner, daftar cocok, wawancara, pengamatan, dan riwayat hidup, sedangkan teknik tes berdasarkan segi kegunaan untuk mengukur siswa diantaranya tes diagnostic, tes formatif, dan tes sumatif<sup>1</sup>, Tes yang baik dapat diketahui dengan melakukan analisis butir soal tes. Informasi yang diperoleh dari kegiatan tersebut meliputi tingkat kesukaran, daya beda, dan efektifitas pengecoh<sup>2</sup>.

Menurut Daryanto<sup>3</sup> dan Ishak<sup>4</sup> bahwa Analisis butir soal merupakan suatu proses untuk mengkaji kualitas pada setiap butir soal. Analisis butir soal bertujuan untuk mengidentifikasi soal yang baik, kurang baik dan soal yang jelek, yang secara langsung memberi isyarat untuk melakukan perbaikan terhadap soal yang teridentifikasi kurang baik.

Idealnya dalam menganalisis butir soal sebagaimana disebutkan di atas, bahwa analisis butir soal meliputi tingkat kesukaran, daya pembeda, dan efektifitas pengecoh dari soal. Sehingga akan menentukan apakah soal tersebut dapat dikatakan bagus atau jelek sehingga dapat memberikan saran untuk merevisi atau mengganti soal tersebut.

<sup>1</sup> Arikunto, Suharsimi. 2010. *Dasar-dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara, hlm. 26&33.

<sup>2</sup> Kashyap, Surekha. *Item Analysis of Multiple Choice Questions. International journal of current research* 7 (2018).

<sup>3</sup> Daryanto.2008. *Evaluasi Pendidikan*. Jakarta : PT Rineka Cipta. hlm.179.

<sup>4</sup> Ishak, Baego dan Syamsuduha. 2011. *Evaluasi Pendidikan*. Makasar. Alauddin press, hlm. 130

Setelah sebuah tes telah terlaksana, sangat penting untuk setiap pendidik agar dapat menganalisis tes yang telah diujikan, agar dapat diketahui bagaimana kualitas tes tersebut.

Berdasarkan hasil observasi penulis, bahwa guru Bahasa Arab di MTs Al-Musyawahar Lembang belum menganalisis hasil UAS mata pelajaran Bahasa Arab tahun pelajaran 2018/2019. Hasil dari ujian siswa belum dijadikan sebagai gambaran apakah soal tersebut sukar, mudah atau sedang sehingga tidak dapat membedakan siswa yang paham dan kurang paham terhadap materi pembelajaran. Maka perlu dilakukan analisis terhadap soal UAS agar dapat diketahui kualitas tes, kemudian hasil tersebut dapat menjadi rujukan dalam pembelajaran bahasa arab juga pembuatan soal berikutnya. Juga perlunya dilakukan penganalisisan terhadap soal tersebut, apakah soal tersebut telah melatih kemampuan berpikir siswa atau justru sebaliknya yang mana hal itu dapat dilihat dari HOTS (Higher Order Thinking Skill) dan LOTS (Lower Order Thinking Skill).

## TINJAUAN PUSTAKA

### Evaluasi

Secara harfiah, kata evaluasi berasal dari bahasa Inggris “evaluation” dalam bahasa arab “*al-taqdir*”, dalam bahasa indonesia berarti “penilaian”. Akar katanya adalah “*value*” dalam bahasa arab “*al-qimah*”, dalam bahasa indonesia berarti “nilai”. jadi secara harfiah, evaluasi pendidikan (*educational evaluation = al-taqdir al-tarbawi*) dapat diartikan sebagai penilaian dalam (bidang) pendidikan atau penilaian mengenai hal-hal yang berkaitan dengan kegiatan pendidikan.<sup>5</sup>

Menurut Ralph Tyler dalam Suharsimi<sup>6</sup> dan Sauri<sup>7</sup> bahwa proses evaluasi merupakan suatu proses untuk menentukan nilai dan mengetahui sejauhmana proses pendidikan telah tercapai dan apakah hasil tersebut telah sesuai dengan apa yang diharapkan. Namun, menurut Cronbach dan Stufflebeam dalam Suharsimi<sup>8</sup> bahwa proses evaluasi bukan sekedar mengukur sejauhmana tujuan telah tercapai, tetapi digunakan juga untuk membuat keputusan. Adapun menurut Subali<sup>9</sup> evaluasi merupakan suatu proses yang sistematis yang dilaksanakan untuk mengetahui tingkat keberhasilan dan efisiensi dari program yang bersangkutan.

Evaluasi dalam pendidikan memiliki peranan penting bahkan mengandung pula makna yang besar bagi perbaikan dan peningkatan mutu pendidikan dan pembelajaran. Ia merupakan barometer untuk mengukur bagaimana keadaan suatu situasi pendidikan

<sup>5</sup> Sudijono, Anas. 2015. *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers. Hlm. 1.

<sup>6</sup> Arikunto, Suharsimi. 2010. *Dasar-dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara, hlm. 3.

<sup>7</sup> Sauri, Sofyan. 2006. *Kesantunan Berbahasa*. Bandung : Rizqi Press, hlm. 158.

<sup>8</sup> Arikunto, Suharsimi. 2010. *Dasar-dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara, hlm. 3.

<sup>9</sup> Subali, Bambang. 2012. *Prinsip Asesmen & Evaluasi Pembelajaran*. Yogyakarta: UNY press, hlm. 1.

atau pembelajaran di suatu kelas atau sekolah, apakah baik atau buruk, memuaskan atau tidak, mengalami kemajuan, kemunduran, atau bahkan kemacetan.<sup>10</sup>

## Tes

Menurut Sudijono<sup>11</sup> dan Matsna bahwa secara harfiah, kata tes berasal dari bahasa perancis kuno “testum” dengan arti “piring untuk menyisihkan logam-logam mulia” (maksudnya dengan menggunakan alat berupa piring itu akan dapat diperoleh jenis-jenis logam mulia yang nilainya sangat tinggi), dalam bahasa inggris ditulis dengan “test” yang dalam bahasa indonesia diterjemahkan dengan “tes”, “tujuan” atau “percobaan” dalam bahasa arab “imtihan”

Menurut Oller dalam Syihabuddin<sup>12</sup> bahwa dalam konteks tes bahasa, mendefinisikan tes bahasa sebagai alat yang digunakan untuk menilai seberapa besar seseorang telah belajar bahasa tersebut. Sedangkan menurut Thu’imah dalam Matsna dan Erta<sup>13</sup> bahwa tes bahasa dipahami sebagai sejumlah pertanyaan atau tugas yang harus dijawab atau direspon siswa dengan tujuan mengukur tingkat mereka dalam penguasaan keterampilan bahasa tertentu.

## Kaidah Penulisan Tes Pilihan Ganda

Soal pilihan ganda adalah soal yang jawabannya harus dipilih dari beberapa kemungkinan jawaban yang telah disediakan, soal tes pilihan ganda dapat digunakan mengukur hasil belajar yang lebih kompleks dan berkenaan dengan aspek ingatan, pengertian, aplikasi, analisis, sintesis dan evaluasi.

Menurut Hazraini dalam menulis soal pilihan ganda harus memperhatikan kaidah-kaidah sebagai berikut:

1. Materi : 1) Soal harus sesuai dengan indikator, 2) Pilihan jawaban harus homogen dan logis ditinjau dari segi materi, 3) Setiap soal harus mempunyai satu jawaban yang benar atau yang paling benar.
2. Kontruksi
  - a. Pokok soal harus dirumuskan secara jelas dan tegas.
  - b. Rumusan pokok soal dan pilihan jawaban harus merupakan pernyataan yang diperlukan saja.
  - c. Pokok soal jangan memberi petunjuk kearah jawaban benar.
  - d. Pokok soal jangan mengandung pernyataan yang bersifat negatif ganda.
  - e. Panjang rumusan pilihan jawaban harus relatif sama.

<sup>10</sup> Saputra, Hatta. 2016. *Pengembangan Mutu Pendidikan Menuju Era Global: Penguatan Mutu Pembelajaran dengan Penerapan HOTS (High Order Thinking Skills)*. Bandung: SMILE's Publishing, hlm. 9.

<sup>11</sup> Sudijono, Anas. 2015. *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers, hlm. 66.

<sup>12</sup> Ishak, Baego dan Syamsuduha. 2011. *Evaluasi Pendidikan*. Makasar. Alauddin press, hlm. 13.

<sup>13</sup> Haig, Yvonne. (2014). *Higher Order Thinking And Assessment. International Seminar on current issues in Primary Education: Prodi PGSD Universitas Muhammadiyah Makasar*, hlm. 43.

- f. Pilihan jawaban jangan mengandung pernyataan, “semua pilihan diatas salah”, atau “semua pilihan jawaban diatas benar”.
  - g. Pilihan jawaban yang berbentuk angka atau waktu harus disusun berdasarkan urutan besar kesilnya nilai angka tersebut, atau kronologinya.
  - h. Gambar, grafik tabel, diagram, dan sejenisnya yang terdapat pada soal harus jelas dan berfungsi.
  - i. Butir soal jangan bergantung pada jawaban soal sebelumnya.
3. Bahasa
- a. Setiap soal harus menggunakan bahasa yang sesuai dengan kaidah Bahasa Indonesia.
  - b. Jangan menggunakan bahasa yang berlaku setempat, jika soal digunakan untuk daerah lain atau nasional.
  - c. Setiap soal harus menggunakan bahasa yang komunikatif.
  - d. Pilihan jawaban jangan mengulang kata atau frase yang bukan merupakan satu kesatuan pengertian.<sup>14</sup>

### ***Higher Order Thinking (HOTS) & Lower Order Thinking Skill (LOTS)***

HOTS atau keterampilan berfikir tingkat tinggi didefinisikan sebagai penggunaan pikiran secara lebih luas untuk menemukan tantangan baru. Kemampuan berfikir tingkat tinggi ini menghendaki seseorang untuk menerapkan informasi baru atau pengetahuan sebelumnya dan memanipulasi informasi untuk menjangkau kemungkinan jawaban dalam situasi baru. Keterampilan berfikir sangat penting dalam proses pendidikan. Orang berfikir dapat mempengaruhi kemampuan belajar, kecepatan dan efektivitas belajar. Oleh karena itu, keterampilan berfikir ini dilakukan dengan proses belajar.

Berfikir tingkat tinggi merupakan jenis pemikiran yang mencoba mengeksplorasi pertanyaan-pertanyaan mengenai pengetahuan yang ada terkait isu-isu yang tidak didefinisikan dengan jelas dan tidak memiliki jawaban yang pasti. Tujuan utama dari HOTS adalah bagaimana meningkatkan kemampuan peserta didik pada level yang lebih tinggi, terutama yang berkaitan dengan kemampuan untuk berfikir secara kritis dalam menerima berbagai jenis informasi, berfikir kreatif dalam memecahkan masalah menggunakan pengetahuan yang dimiliki serta membuat keputusan dalam situasi yang kompleks. Direktorat Jenderal Guru dan Tenaga Kependidikan Kementerian Pendidikan dan Kebudayaan 2018 telah mengklasifikasikan proses kognitif antara HOTS & LOTS, yaitu sebagai berikut :

**Tabel 1 Kriteria Kognitif HOTS & LOTS**

Proses Kognitif			Definisi
C1	L	Mengingat	Mengambil pengetahuan yang relevan dari ingatan

<sup>14</sup> Hazraini. 2017. *Upaya Meningkatkan Kompetensi Guru Kelas Dalam Penyusunan Soal Pilihan Ganda Yang Baik dan Benar Melalui Pendampingan Berbasis KKG Semester Satu Tahun Pelajaran 2017/2018 di SD Negeri 40 Cakranegara. JUPE, Vol. 2 No.2, hlm, 17.*

C2	O	Memahami	Membangun arti dari proses pembelajaran, termasuk komunikasi lisan, tertulis, dan gambar
C3	S	Menerapkan/ mengaplikasikan	Melakukan atau menggunakan prosedur di dalam situasi yang tidak biasa.
C4	H O T S	Menganalisis	Memecah materi ke dalam bagian-bagiannya dan menentukan bagaimana bagian-bagian itu terhubung antarbagian dan ke struktur atau tujuan keseluruhan
C5		Menilai/ mengevaluasi	Membuat pertimbangan berdasarkan kriteria atau standar
C6		Mengkreasi/mencipta	Menempatkan unsur-unsur, secara bersama-sama untuk membentuk keseluruhan secara koheren atau fungsional, menyusun kembali unsur-unsur ke dalam pola atau struktur baru

(Kemendikbud, 2018)

## METODOLOGI

Metode yang digunakan dalam penelitian ini adalah deskriptif. Populasi dalam penelitian ini difokuskan terhadap seluruh kelas VIII MTs Al-Musyawahar Lembang yang berjumlah 185 orang. Adapun sampel dalam penelitian ini adalah 30 orang dari jumlah keseluruhan, diambil dari kelas E dan F menggunakan nonprobability sampling dengan teknik sampling kuota.

Instrumen pada penelitian ini adalah dokumentasi yang digunakan untuk mengumpulkan data berupa kisi-kisi penilaian akhir semester (PAS), soal tes beserta kunci jawabannya, dan lembar jawaban siswa untuk tes UAS Bahasa Arab kelas VIII MTs Al-Musyawahar Lembang yang telah ada, sebanyak 40 butir soal dan 30 lembar jawaban siswa yang akan dianalisis kualitasnya dari segi validitas, reliabilitas, tingkat kesukaran, daya pembeda, efektifitas pengecoh dan menganalisis soal berbasis, apakah termasuk *Hots* atau *Lots*. Penelitian dibantu dengan aplikasi software Anates versi 4.

## HASIL DAN PEMBAHASAN

### Validitas dan Reliabilitas Tes

Berdasarkan hasil analisis terhadap 40 butir soal Ulangan Akhir Semester Bahasa Arab menggunakan *software* anates versi 4 diperoleh hasil bahwa terdapat soal yang masuk dalam kategori valid/signifikan berjumlah 7 butir soal (17,5%) 24 butir soal(60%) sangat valid/sangat signifikan dan 9 (22,5)% butir soal masuk ke dalam kategori tidak valid. Butir soal yang tidak valid sebaiknya dilakukan perbaikan ulang/revisi dan butir soal yang valid dapat digunakan kembali untuk tes. Hal ini sesuai

dengan teori validitas menurut Anas Sudijono bahwa butir soal yang memiliki validitas tinggi mencerminkan soal tersebut telah memiliki kehandalan dan tidak perlu diragukan lagi ketepatannya dalam mengukur kemampuan peserta didik. Untuk butir soal yang memiliki validitas rendah mencerminkan soal tersebut tidak valid sehingga perlu dilakukan tindakan terhadap soal tersebut.

Sementara hasil analisis reliabilitas menunjukkan bahwa soal UAS memiliki reliabilitas sebesar 0,89. Dengan koefisien reliabilitas sebesar 0,89 menunjukkan bahwa tes yang digunakan sudah memiliki reliabilitas yang tinggi. Hal ini berdasarkan pada kriteria reliabilitas menurut Masidjo bahwa koefisien korelasi 0,71-0,90 termasuk dalam kualifikasi tinggi.

### **Kualitas Tes**

#### **a. Segi Kaidah**

Terdapat 27 (67,5%) butir soal yang tidak memenuhi kriteria penulisan tes pilihan ganda, dengan nomor soal 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 21, 22, 23, 25, 27, 28, 30, 31, 34, 35. Hal itu menunjukkan bahwa soal termasuk kategori sedang, berdasarkan tabel berikut:

**Tabel 2 Kriteria Keberhasilan Proses Pembelajaran Siswa Dan Guru Dalam %**

No.	Tingkat Keberhasilan	Predikat Keberhasilan
1	86-100 %	Sangat Tinggi
2	71-85 %	Tinggi
3	56-70 %	Sedang
4	41-55 %	Rendah
5	<40 %	Sangat rendah
	Rentang 15 %	

(Sumber : adaptasi dari Agip dkk, 2009 hlm 41)

#### **b. Segi Butir Soal**

##### **Tingkat Kesukaran**

Hasil analisis tingkat kesukaran soal yang dilakukan terhadap 40 butir soal pilihan ganda menggunakan *software* Anates versi 4 diperoleh hasil bahwa 5 butir soal masuk dalam kategori sangat mudah (12,5%), 3 butir soal dalam kategori mudah (7,5%), 23 butir soal dalam kategori sedang (57,5%), 7 butir soal dalam kategori sukar (17,5%), dan 2 butir soal dalam kategori sangat sukar (5%). Berdasarkan perhitungan tersebut, terdapat beberapa soal yang termasuk kategori sangat mudah dan sangat sukar. Hal demikian tidak baik untuk dijadikan sebuah tes, karena jika soal terlalu sulit, akan membuat siswa menjadi putus asa dikarenakan ia merasa tidak mampu untuk menjawab soal tersebut. Untuk soal yang terlalu mudah, hal itu akan membuat siswa tidak bersungguh-sungguh atau tidak merangsang siswa dalam meningkatkan usahanya dalam memecahkan masalah.



### **Daya Pembeda**

Daya pembeda soal berkaitan dengan kemampuan soal untuk mengetahui dan membedakan antara siswa yang pandai (menguasai materi) dan siswa yang kurang pandai (tidak/kurang menguasai materi) nilai daya beda dari hasil perhitungan di atas, diperoleh hasil bahwa 22 butir soal termasuk kategori sangat memuaskan, 3 butir soal memuaskan, 8 butir soal tidak memuaskan dan 7 butir soal termasuk kategori sangat tidak memuaskan.

Oleh karena itu, dari hasil penelitian dapat 3 kemungkinan tindak lanjut. Yaitu: 1. Soal dengan kategori memuaskan dan sangat memuaskan dapat diterima, disimpan di bank soal dan dapat digunakan lagi pada tes yang akan datang karena kualitasnya memadai. 2. Daya beda yang mempunyai kategori tidak memuaskan dapat ditolak/dibuang karena tidak dapat digunakan lagi untuk mengevaluasi hasil siswa juga dapat direvisi lagi sehingga bisa digunakan pada tes yang akan datang. 3. Soal dengan daya beda berkategori sangat tidak memuaskan dibuang dan tidak boleh digunakan pada tes yang akan datang.

### **Efektifitas Pengecoh**

Sama halnya dengan daya pembeda, bahwa efektifitas pengecohpun harus membedakan jawaban antara siswa kelompok atas dan siswa kelompok bawah. Berdasarkan analisis yang telah dilakukan melalui *software* Anates versi 4 bahwa efektifitas pengecoh soal UAS Bahasa Arab dapat dikatakan cukup. Sebanyak 15 butir soal mempunyai pengecoh yang tidak baik/tidak berfungsi, dan 25 butir soal dikatakan baik. Namun, meskipun angka efektifitas pengecoh menunjukkan lebih unggul, akan tetapi kualitas efektifitas pengecoh yang tidak baik memiliki/masuk dalam kategori sangat buruk. Sehingga harus terus diperbaiki dan dibenahi menjadi lebih baik.

### **HOTS**

Berdasarkan hasil analisis tingkat kognitif terhadap soal UAS Bahasa Arab kelas VIII MTs Al-Musyawahrah Lembang tahun ajaran 2018-2019 yang dianalisis berdasarkan kategori HOTS & LOTS Kemendikbud menghasilkan sebuah kesimpulan bahwa keseluruhan soal merupakan soal berkategori LOTS, dengan angka C1 sebanyak 2 butir soal, C2 2 butir soal, dan C3 sebanyak 36 butir soal. Hal ini menandakan bahwa soal tidak mengukur siswa untuk memiliki kemampuan berfikir tinggi/HOTS.

## **KESIMPULAN**



Setelah melakukan analisis terhadap kualitas butir tes UAS Bahasa Arab kelas VII MTs Al-Musyawaharah Tahun Ajaran 2018-2019, dapat ditarik kesimpulan sebagai berikut:

Pertama : ditinjau dari segi validitas tes sebanyak 7 butir soal (17,5%) termasuk valid/signifikan, 24 butir soal(60%) sangat valid/sangat signifikan dan 9 (22,5)% butir soal masuk ke dalam kategori tidak valid. Sementara hasil analisis reliabilitas menunjukkan bahwa soal UAS memiliki reliabilitas sebesar 0,89. Dengan koefisien reliabilitas ( $r_{11}$ ) sebesar 0,89 menunjukkan bahwa tes yang digunakan sudah memiliki reliabilitas yang tinggi.

Kedua : dilihat dari analisis kualitatif terdapat 25 (62,5%) butir soal yang tidak memenuhi kriteria penulisan tes pilihan ganda, dengan nomor soal 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 21, 22, 23, 25, 27, 28, 30, 31, 34, 35 Hal itu meunjukkan bahwa soal termasuk kategori sedang. Dilihat dari analisis kuantitatif dari segi tingkat kesukaran diperoleh hasil bahwa 5 butir soal masuk dalam kategori sangat mudah (12,5%), 3 butir soal dalam kategori mudah (7,5%), 23 butir soal dalam kategori sedang (57,5%), 7 butir soal dalam kategori sukar (17,5%), dan 2 butir soal dalam kategori sangat sukar (5%) dengan kesimpulan bahwa tingkat kesukaran belum proporsional. Dari segi daya pembeda diperoleh hasil bahwa 22 butir soal termasuk kategori sangat memuaskan, 3 butir soal memuaskan, 8 butir soal tidak memuaskan dan 7 butir soal termasuk kategori sangat tidak memuaskan, dengan kesimpulan bahwa daya pembeda tergolong sedang. Dari segi efektifitas pengecoh sebanyak 19 butir soal mempunyai pengecoh yang tidak baik/tidak berfungsi, dan 21 butir soal dikatakan baik hal ini menunjukkan efektifitas pengecoh dikatakan berfungsi dengan cukup.

Ketiga : ditinjau dari segi HOTS & LOTS, bahwa keseluruhan soal merupakan soal berkategori LOTS, dengan angka C1 sebanyak 2 butir soal, C2 2 butir soal, dan C3 sebanyak 36 butir soal.

#### DAFTAR PUSTAKA

- Agip, Zainal dkk, 2009. *Penelitian Tindakan Kelas untuk Guru SD, SLB, dan TK*. Bandung : Yrama Widya.
- Arifin, Zainal. 2013. *Evaluasi Pembelajaran (Prinsip, Teknik, Prosedur)*. Bandung: PT. Remaja Rosdakarya.
- Arikunto, Suharsimi. 2010. *Dasar-dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara.
- Daryanto.2008. *Evaluasi Pendidikan*. Jakarta : PT Rineka Cipta.

Direktorat Jenderal Guru dan Tenaga Kependidikan Kementerian Pendidikan dan Kebudayaan 2018.

Haig, Yvonne. (2014). *Higher Order Thinking And Assessment. International Seminar on current issues in Primary Education*: Prodi PGSD Universitas Muhammadiyah Makasar.

Hazraini. 2017. *Upaya Meningkatkan Kompetensi Guru Kelas Dalam Penyusunan Soal Pilihan Ganda Yang Baik dan Benar Melalui Pendampingan Berbasis KKG Semester Satu Tahun Pelajaran 2017/2018 di SD Negeri 40 Cakranegara. JUPE, Vol. 2 No.2.*

Heong, Y.M., dkk. 2011. *The level of Marzano Higher Order Thinking Skills Among Trchnical Education Students. International Jurnal of Social and Humanity, Vol.1, No.2, July 2011, 121-12.*

Ishak, Baego dan Syamsuduha. 2011. *Evaluasi Pendidikan*. Makasar. Alauddin press

Kashyap, Surekha. *Item Analysis of Multiple Choice Questions. International journal of current research 7 (2018) : hlm..*

Saputra, Hatta. 2016. *Pengembangan Mutu Pendidikan Menuju Era Global: Penguatan Mutu Pembelajaran dengan Penerapan HOTS (High Order Thinking Skills)*. Bandung: SMILE's Publishing.

Sauri, Sofyan. 2006. *Kesantunan Berbahasa*. Bandung : Rizqi Press.

Subali, Bambang. 2012. *Prinsip Asesmen & Evaluasi Pembelajaran*. Yogyakarta:UNY press.

Sudijono, Anas. 2015. *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers.