

## PENGEMBANGAN INSTRUMEN PENILAIAN KOGNITIF BERBASIS GOOGLE FORM PADA MATA PELAJARAN BIOLOGI

**Umni Alwiyah**

Universitas Islam Negeri Alauddin Makassar  
Email: ualwiyah@gmail.com

**St. Syamsudduha**

Universitas Islam Negeri Alauddin Makassar  
Email: st.syamsudduha@uin-alauddin.ac.id

**Ahmad Ali**

Universitas Islam Negeri Alauddin Makassar  
Email: ahmad.ali@uin-alauddin.ac.id

### **Abstrak**

Penelitian ini bertujuan untuk mengembangkan instrumen penilaian kognitif berbasis *Google Form* yang valid, reliabel, daya pembeda yang baik, tingkat kesukaran yang bervariasi (mudah, sedang, dan sukar), dan efektivitas pengecoh yang bekerja pada Mata Pelajaran Biologi Kelas X SMA Negeri 1 Takalar. Instrumen penilaian kognitif yang dikembangkan berbentuk pilihan ganda yang terformat dalam *Google Form*. Jenis penelitian ini adalah penelitian pengembangan tipe *Formative Research* yang dikembangkan oleh Tessmer, pengembangan ini terdiri dari 3 tahap yaitu *Preliminary, Self Evaluation dan Prototyping*. Berdasarkan hasil penelitian maka diperoleh koefisien validitas isi yaitu 1 yang berada pada kriteria tinggi dan 50% dari 30 butir soal dalam kriteria valid. Hasil uji reliabilitas menggunakan rumus KR-20 yang diperoleh adalah 0,70 dengan kriteria tinggi. Daya pembeda sebesar 70% dengan 20% termasuk kriteria cukup, 30% termasuk kriteria baik, dan 20% termasuk kriteria sangat baik. Tingkat kesukaran yang bervariasi yakni mudah, sedang, dan sukar, masing-masing menunjukkan perbandingan 23%:54%:23%. Efektivitas pengecoh yang bekerja dengan baik dengan persentase 80%, yakni 33% termasuk kriteria cukup dan 47% termasuk kriteria efektif.

**Kata Kunci:** penilaian kognitif, analisis butir soal, *google form*

### **Abstract**

*This research aims to develop Google Form-based cognitive assessment instruments and to assess its validity, reliability, level of difficulty, and distractors' distribution. The subjects of this research were the IPA students of Class X in SMAN 1 Takalar. The Cognitive assessment instrument used was in the format of multiple-choice in Google Form. This research development was adapt Tessmer model that consisted of 3 stages (preliminary, self-evaluation, and prototyping). The results showed that the coefficient validity was 1 with high criteria, and there were 50% of 30 items in the valid criteria. The reliability of the test using the KR-20 formula obtained 0,70 with high criteria. The discrimination power showed 70% where the criteria were 20% was sufficient, 30% was good, and 20% very good. The levels of difficulty of the test were easy, medium, and difficult with the ratio of a value of each level were 23%:54%:23%. The*

*effectiveness of distractors' distribution has a percentage of 80%, with 33% was in enough criteria and 47% was effective criteria.*

**Keywords:** *cognitive assessment, item analysis, google form*

## **PENDAHULUAN**

Saat ini guru menghadapi peserta didik yang berbeda jauh dengan masa sebelumnya. Peserta didik sekarang lahir di zaman yang ekstra modern. Teknologi yang sangat berkembang pesat memudahkan peserta didik mengakses informasi dengan mudah dan cepat. Penggunaan teknologi dengan memanfaatkan jaringan internet telah banyak digunakan oleh guru dalam proses pembelajarannya. Proses pembelajaran melibatkan interaksi timbal balik antara guru dan peserta didik yang bersifat edukatif (Purwati & Nugroho, 2018). Setelah proses pembelajaran, kegiatan selanjutnya adalah melakukan evaluasi.

Evaluasi adalah salah satu cara untuk menilai suatu objek yang memerlukan data hasil pengukuran dari berbagai dimensi misalnya pengetahuan, sikap, keterampilan dan sebagainya. Penilaian suatu objek membutuhkan adanya instrumen penilaian, baik untuk menilai aspek kognitif, afektif, maupun psikomotorik peserta didik. Instrumen penilaian tersebut harus memenuhi syarat yakni valid dan reliabel (Hidayat, 2017). Teori kognitif dari Benjamin S. Bloom telah direvisi oleh D. R. Anderson Krathwohl, dimana dimensi kognitif berhubungan dengan enam tingkatan atau kategori yang disajikan dalam verba mengingat, memahami, menerapkan, menganalisis, mengevaluasi, dan mencipta (Paidi, Djukri, Yulaikah, & Alfindasari, 2017).

Evaluasi yang diberikan kepada peserta didik saat ini masih banyak menggunakan kertas, seperti pemberian ulangan dan tugas rumah. Pemberian waktu pengerjaan ulangan, terkadang tidak berjalan sesuai rencana waktu pengerjaan setiap butir soal. Selain itu, lamanya waktu pemeriksaan lembar jawaban oleh guru dan proses penginputan nilai. Hal seperti inilah yang sangat membutuhkan banyak kertas, seperti kita ketahui, secara umum bahan pembuatan kertas adalah kayu (Rahayu & Listiyadi, 2014).

Kebutuhan akan kertas yang selalu meningkat, sehingga memungkinkan terjadinya eksploitasi hutan juga akan meningkat. Pada tahun 2013, Badan Penelitian dan Pengembangan Industri (BPPI) merilis hasil penelitian yang menyatakan bahwa Indonesia mampu memproduksi pulp dan kertas berjuta-juta ton. Hal ini menjadikan Indonesia menduduki peringkat ke-9 untuk produsen *pulp* terbesar di dunia dan

peringkat ke-6 untuk produsen kertas terbesar di dunia. Bahan pembuatan kertas saat ini masih didominasi oleh kayu, untuk itu Indonesia harus mencari alternatif lain agar dapat mengurangi eksploitasi hutan (Rosmainar, 2017).

Adanya teknologi canggih seperti sekarang ini, diharapkan mampu mengurangi penggunaan kertas yang berlebihan. Untuk meminimalisir hal tersebut, Indonesia telah menggunakan sistem evaluasi yang tidak lagi menggunakan kertas (Paper Based Test) tetapi beralih menggunakan komputer (Computer Based Test), sistem ini sudah berlaku di beberapa satuan tingkat pendidikan di Indonesia. Sistem ini juga cukup efisien karena memudahkan guru untuk menilai hasil pekerjaan peserta didik. Sistem CBT ini juga digunakan dalam seleksi nasional. Pemerintah Indonesia telah mengaplikasikan alat evaluasi pembelajaran CBT dalam sistem ujian nasional, yang disebut UNBK (Ujian Nasional Berbasis Komputer) yang dilakukan secara *online* menggunakan komputer dan untuk mengaksesnya ke server pusat dibutuhkan jaringan internet yang memadai (Purwati & Nugroho, 2018).

Konsep evaluasi pembelajaran berbasis teknologi ini juga bisa diterapkan dalam evaluasi pembelajaran harian di kelas dibantu dengan penggunaan instrumen berbasis teknologi juga. Salah satu aplikasi yang dapat dimanfaatkan sebagai instrumen penilaian evaluasi berbasis *online* adalah *Google Form*, aplikasi ini berupa formulir yang merupakan bagian dari penyimpanan *Google Drive*. Syarat untuk menggunakannya hanya dengan memiliki akun Google saja bagi pembuat *Form* (Mardiana & Purnanto, 2017). Dibandingkan dengan menggunakan CBT, *Google Form* lebih praktis karena dapat diakses di *android/smartphone* yang mudah untuk dibawa kemana-mana, dan mengaksesnya dapat dimana saja selama jaringan internet terjangkau.

Kelebihan dari *Google Form* juga karena merupakan layanan gratis dari Google yang menyediakan sistem survei *online* yang mendukung berbagai jenis pertanyaan seperti teks, paragraf, pilihan ganda, memilih dari daftar, kotak centang, skala, dan jaringan. Jenis teksnya digunakan untuk pertanyaan yang jawabannya singkat dan jawaban yang panjang atau berupa paragraf (Kim & Park, 2012). Proses pembuatan template *Google Form* di dalamnya terdapat berbagai macam karakteristik pilihan yang diberikan. Seperti background berbagai kategori, bentuk font, menambahkan gambar dan video, dan lain-lain, yang dapat disesuaikan dengan kebutuhan (Iqbal, et al, 2018).

Aplikasi *Google Form* dapat pula digunakan setiap orang untuk membuat kuis ataupun kuesioner *online* menggunakan laptop, komputer atau android yang terhubung dengan internet (Batubara & Ariani, 2016). Kemudian bisa dihubungkan dengan proyektor, sehingga peserta didik dapat melihat jawaban yang mereka kirimkan, baik itu jawaban mereka sendiri atau jawaban rekan-rekannya (Djenno, M., Insua, G.M. & Pho, A, 2015). Selain memiliki banyak kelebihan, *Google Form* juga memiliki kelemahan, seperti belum ada pilihan jenis pertanyaan yang akan disediakan dalam tabel dan tidak ada pilihan untuk huruf miring untuk peserta didik menjawab pertanyaan tertentu (Rodriguez, 2018). Penggunaan *Google Form* dapat memudahkan guru dalam pemberian evaluasi kepada peserta didik, termasuk memudahkan penilaian hasil evaluasi, karena nilai yang peserta didik peroleh dapat dilihat langsung dalam *template*.

Berdasarkan hasil wawancara peneliti bersama salah seorang guru Biologi kelas X SMA Negeri 1 Takalar, diketahui bahwa di sekolah ini memiliki sarana dan prasarana yang lengkap seperti laboratorium komputer beserta wifi, semua peserta didik juga diperkenankan membawa smartphone atau android ke sekolah dan digunakan dalam pembelajaran jika mencari informasi mengenai materi yang dipelajari. Instrumen yang digunakan untuk evaluasi peserta didik seperti tugas harian, ulangan harian, MID semester, dan juga Ujian Akhir Semester masih menggunakan kertas, dengan pemeriksaan lembar jawaban secara manual. Adanya pandemik Covid-19 memaksa pemerintah mengeluarkan kebijakan belajar dalam jaringan (*online*) memaksa guru harus merubah sistem dan strategi pembelajaran termasuk dalam proses evaluasi hasil belajar peserta didik. Oleh karena itu, sangat penting dilakukan pengembangan instrumen penilaian berbasis teknologi, yang berkualitas dan memenuhi standar instrumen yang baik. Perangkat penilaian berbasis teknologi ini diharapkan dapat menjadi solusi bagi guru dalam melakukan evaluasi serta mengurangi penggunaan kertas yang berlebihan, terkhusus pada penilaian kognitif peserta didik.

## **METODE PENELITIAN**

Penelitian ini merupakan penelitian dan pengembangan dengan jenis *Formative Research* (Tessmer 1993). Pengembangan ini terdiri atas 3 tahap yaitu tahap pertama *Preliminary*, tahap kedua *Self Evaluation* yang meliputi analisis dan desain, tahap ketiga *Prototyping* yang meliputi *expert reviews*, *one-to-one*, dan *small group* serta *field test*. Instrumen yang peneliti kembangkan adalah instrumen penilaian kognitif sub materi

Invertebrata berbasis *Google Form*, yang valid, reliabel, memiliki daya pembeda baik, tingkat kesukaran yang bervariasi, dan efektivitas pengecoh yang bekerja. Lokasi uji coba di SMA Negeri 1 Takalar. Berikut ini beberapa tahapan pelaksanaan penelitian yang dilakukan peneliti:

### **Tahap Preliminary**

Persiapan awal yang dilakukan adalah pengkajian referensi dari beberapa sumber terkait penelitian. Setelah mendapatkan beberapa informasi dan teori yang cukup, selanjutnya melakukan koordinasi dengan sekolah tempat penelitian.

### **Tahap Self Evaluation**

Tahap ini terbagi menjadi dua, yakni analisis dan desain. Tahap analisis yang dilakukan adalah analisis kurikulum, analisis peserta didik, dan analisis materi. Tahap selanjutnya adalah mendesain instrumen yang dikembangkan. Desain instrumen yang dikembangkan, disesuaikan dengan indikator dan tujuan pembelajaran. Desain instrumen tersebut berupa desain kisi-kisi soal dan instrumen tes dalam bentuk pilihan ganda dengan kunci jawabannya sebanyak 30 butir soal. Produk ini kemudian dimasukkan dalam *Google Form*, hasil desain disebut *Prototype I*.

### **Tahap Prototyping**

#### *Expert Review (Pakar)*

Produk yang telah didesain, kemudian dinilai dan dievaluasi oleh dua orang pakar dalam bidang materi Biologi dan pendidikan Biologi. Pakar bertindak sebagai validator menelaah materi, konstruk, dan bahasa pada produk yang telah didesain tersebut. Semua saran dan tanggapan dari validator dituliskan pada lembar validasi sebagai bahan untuk merevisi *Prototype* yang telah dibuat.

#### *One-to-one*

Tahap selanjutnya, peserta didik yang terpilih menjadi non subjek uji coba adalah peserta didik yang berkemampuan tinggi dan sebaya dengan subjek uji coba. Hasil tes dan komentar peserta didik sebagai subjek uji coba menjadi acuan merevisi produk, hasil dari revisi tersebut disebut *Prototype II*.

#### *Small Group*

Hasil revisi yang dilakukan dari tahap *Expert Review* dan *One-to-one*, dijadikan dasar untuk merevisi *Prototype I* menjadi desain *Prototype II*. Kemudian diujicobakan pada *Small Group* yakni enam orang peserta didik non subjek uji coba yang terdiri dari

dua peserta didik berkemampuan tinggi, dua peserta didik berkemampuan sedang, dan dua peserta didik lainnya berkemampuan rendah pada mata pelajaran Biologi. Hasil tes dan komentar keenam peserta didik, menjadi acuan produk direvisi kembali. Hasil revisi disebut *Prototype III*.

*Field Test* (Uji coba lapangan)

Tahap terakhir dari model pengembangan ini adalah *field test*, hasil revisi dari tahap *Small Group* yang menghasilkan *Prototype III*, kemudian diujicobakan pada 35 peserta didik yang menjadi subjek penelitian. Kemudian hasil uji coba tersebut dianalisis, melalui analisis butir soal yang terdiri dari lima kriteria. Hasil dari analisis ini menunjukkan kualitas setiap butir soal yang telah diujicobakan.

*Analisis Validitas Instrumen*

Sebuah instrumen harus melewati persyaratan valid. Validitas instrumen meliputi validitas setiap butir soal dan validitas keseluruhan instrumen. Analisis validitas isi dilakukan dengan memperhatikan relevansi oleh dua orang pakar ahli. Relevansi kedua pakar dihitung menggunakan rumus Gregory (Gregory, 2007 dalam Retnawati, 2016):

$$\text{Validasi isi} = \frac{D}{(A + B + C + D)}$$

Tabel 1. Model Relevansi Validator

		Validator 1	
		Tidak relevan Skor (1-2)	Relevan Skor (3-4)
Validator 2	Tidak relevan Skor (1-2)	A	B
	Relevan Skor (3-4)	C	D

Validitas isi dengan koefisien korelasi  $\geq 0,81$  termasuk kriteria tinggi,  $\leq 0,40 - \leq 0,80$  termasuk kriteria sedang, dan  $\leq 0,40$  termasuk kriteria rendah (Retnawati, 2016). Sedangkan untuk validitas empiris butir soal dapat dihitung menggunakan rumus statistik, yaitu rumus korelasi *product moment pearson* sebagai berikut (Yusuf, 2015):

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\}\{N \sum Y^2 - (\sum Y)^2\}}}$$

Keterangan:

$r_{xy}$  = Koefisien korelasi antara variabel X dengan variabel Y

N = Jumlah responden

$\sum X$  = Jumlah variabel X

$\sum Y$  = Jumlah variabel Y

$\sum XY$  = Jumlah perkalian X dan Y

$\sum X^2$  = Jumlah kuadrat dari sebaran X

$\sum Y^2$  = Jumlah kuadrat dari sebaran Y

Validitas empiris butir soal termasuk kriteria valid jika  $r_{xy} \geq r_{tabel}$  dan termasuk kriteria tidak valid jika  $r_{xy} < r_{tabel}$  (Hamirul, 2020).

#### *Analisis Reliabilitas Instrumen*

Adapun untuk menghitung reliabilitas dari instrumen tes yaitu dihitung dengan menggunakan rumus KR-20 sebagai berikut (Yusuf, 2015):

$$r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{S^2_t - \sum pq}{S^2_t} \right)$$

Keterangan :

$r_{11}$  = Reliabilitas instrumen secara keseluruhan

n = Jumlah butir soal dalam satu instrumen

p = Proporsi subjek yang menjawab butir soal dengan benar

q = Proporsi subjek yang menjawab butir soal dengan salah

$S^2_t$  = Variance dari instrumen (Kuadrat St)

St = Standar deviasi dari skor total

Instrumen dapat dikatakan *reliable* dengan kriteria tinggi jika nilai koefisien korelasi  $r_{11} \geq 0,70$  dan dikatakan *unreliable* dengan kriteria rendah jika  $r_{11} < 0,70$  (Rosidin, 2017).

#### *Analisis Daya Pembeda Instrumen*

Daya pembeda merupakan pengukuran sejauh mana soal mampu membedakan antara peserta didik yang memiliki kemampuan tinggi dan memiliki kemampuan rendah dalam menguasai kompetensi berdasarkan kriteria tertentu. Untuk menghitung daya pembeda setiap butir soal digunakan rumus sebagai berikut (Syamsudduha, 2014):

$$D = \frac{BA}{JA} - \frac{BB}{JB}$$

Keterangan:

D = Indeks diskriminasi

BA = Jumlah kelompok atas yang menjawab benar

BB = Jumlah kelompok bawah yang menjawab benar

JA = Banyaknya peserta kelompok atas

JB = Banyaknya peserta kelompok bawah

Jika nilai  $Dp \leq 0,00$  maka butir soal termasuk kriteria sangat jelek,  $0,00 < Dp \leq 0,20$  termasuk kriteria jelek,  $0,20 < Dp \leq 0,40$  termasuk kriteria cukup,  $0,40 < Dp \leq 0,70$  termasuk kriteria baik, dan jika  $0,70 < Dp \leq 1,00$  maka termasuk kriteria sangat baik (Syamsudduha, 2014).

#### *Analisis Tingkat Kesukaran Instrumen*

Soal yang baik adalah soal yang memiliki tingkat kesukaran yang seimbang atau proporsional (Arifin, 2013). Dalam artian, soal tidak memiliki tingkat kesukaran yang tinggi dan juga tidak terlalu rendah atau mudah dikerjakan. Tingkat kesukaran biasanya diberi simbol P maksud dari kata proporsi yang dilambangkan sebagai Indeks kesukaran. Untuk mendapatkan indeks kesukaran suatu soal, digunakan rumus sebagai berikut (Syamsudduha, 2014).

$$P = \frac{B}{JS}$$

Keterangan:

P = Indeks kesukaran

B = Banyaknya peserta didik yang menjawab butir soal dengan benar

JS = Jumlah seluruh peserta didik

Butir soal termasuk kriteria sangat sukar jika nilai P 0,00 - 0,30, termasuk kriteria sedang jika nilai P 0,31 - 0,70, dan termasuk kriteria mudah jika nilai P 0,71 - 1,00 (Syamsudduha, 2014).

#### *Analisis Keefektifan Pengecoh Instrumen*

Soal bentuk pilihan ganda memiliki opsi atau alternatif jawaban yang merupakan sebuah pengecoh. Butir soal yang baik, pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya soal yang kurang baik, pengecohnya akan dipilih secara tidak merata. Indeks pengecoh dihitung dengan rumus (Arifin, 2013):

$$IP = \frac{P}{(N-B)/(n-1)} \times 100\%$$



Keterangan:

IP = Indeks pengecoh

P = Jumlah peserta didik yang menjawab salah pada setiap soal (terkecoh)

N = Jumlah peserta didik yang ikut tes

B = Jumlah peserta didik yang menjawab benar pada setiap soal

n = Jumlah alternatif jawaban (opsi)

I = Bilangan tetap

Jika nilai Ip lebih dari 200% maka pengecoh termasuk kriteria sangat jelek, jika 0% - 25% atau 176% - 200% termasuk jelek, 26% - 50% atau 151% - 175% termasuk kurang baik, 51% - 75% atau 126% - 150% termasuk baik, dan jika nilai Ip 76% - 125% maka pengecoh termasuk sangat baik (Arifin, 2013).

## HASIL DAN PEMBAHASAN

Instrumen penilaian kognitif berbasis *Google Form* pada sub materi Invertebrata, yang diujikan pada Kelas X IPA 1 SMA Negeri 1 Takalar dengan jenis pengembangan *Formative Research* (Tessmer 1993). Hasil yang diperoleh pada penelitian yang telah dilakukan, sebagai berikut:

### **Tahap Preliminary**

Tahap ini dilakukan dengan mengumpulkan beberapa referensi yang relevan dengan penelitian yang dilakukan, yakni pengembangan instrumen penilaian kognitif dan penggunaan aplikasi *Google Form*. Selain mengumpulkan beberapa referensi, melakukan komunikasi dengan pihak sekolah SMA Negeri 1 Takalar yang merupakan lokasi penelitian.

### **Tahap Self Evaluation**

#### *Analisis*

Tahap analisis terbagi menjadi tiga, yang pertama analisis kurikulum, analisis ini merupakan hal yang pertama untuk diketahui dalam tahapan analisis. Diketahui bahwa kurikulum yang diterapkan adalah K13 (Kurikulum 2013). Kedua analisis peserta didik, terkait subjek uji coba maupun non subjek uji coba. Subjek uji coba adalah peserta didik kelas X IPA 1 berjumlah 35 orang, pada tahap penelitian *field test*. Sedangkan non subjek uji coba adalah peserta didik kelas X IPA 2 berjumlah 7 orang dengan dua tahap penelitian yakni tahap *one-to-one* yang berjumlah 1 orang peserta didik dan tahap *small group* yang berjumlah 6 orang peserta didik. Semua peserta didik di sekolah ini

diperbolehkan membawa *smartphone* dan digunakan selama proses pembelajaran jika diperlukan, sehingga dapat memudahkan penelitian yang menggunakan *smartphone* untuk mengakses aplikasi *Google Form*. Ketiga, analisis materi, terkait materi yang menjadi fokus penelitian dalam pengembangan instrumen tes yakni sub materi Invertebrata dari materi Kingdom Animalia yang diajarkan pada semester genap. Materi tersebut terdiri dari delapan *filum* beserta contoh spesiesnya, hewan-hewan tersebut ada yang sering kita jumpai di lingkungan sekitar dan ada juga yang jarang kita jumpai. Bahkan ada beberapa hewan yang kadang sebagian orang menganggapnya tumbuhan.

### **Desain**

Tahap desain dilakukan berdasarkan hasil analisis sebelumnya. Instrumen tes dirancang berdasarkan materi dan indikator soal yang sesuai dengan kompetensi dasar, kemudian membuat kisi-kisi dan butir soal yang berbentuk pilihan ganda sebagai desain awal. Kisi-kisi soal yang dikembangkan meliputi identitas yang terdiri dari nama sekolah, mata pelajaran, dan kelas, kemudian kompetensi inti dan kompetensi dasar, materi, indikator soal, tingkatan kognitif, bentuk soal, dan nomor soal, serta pedoman penskoran. Tingkatan kognitif yang terdapat dalam kisi-kisi yang telah dikembangkan beracuan pada aspek kognitif taksonomi Bloom revisi D. R. Krathwohl Anderson C2 (memahami), C3 (menerapkan), dan C4 (menganalisis). Sebanyak 30 butir soal yang dikembangkan pada penelitian ini. Soal tersebut diformat dalam bentuk *Google Form*. Hasil pengembangan ini, dijadikan sebagai *Prototype I* yang selanjutnya divalidasi oleh dua pakar ahli.

### **Tahap Prototyping**

#### *Expert Review (Pakar)*

Instrumen yang telah dibuat selanjutnya diberikan kepada masing-masing pakar ahli (validator) yang berjumlah dua orang. Kemudian memberikan penilaian terhadap instrumen, penilaian tersebut dikenal dengan istilah lembar validasi yang meliputi aspek materi, konstruksi, dan bahasa. Berikut ini rekapitulasi hasil relevansi kedua validator, dengan menggunakan rumus validitas isi Gregory (Retnawati, 2016):

Tabel 2. Rekapitulasi Hasil Relevansi Validator

		<b>Validator 1</b>	
		<b>Tidak relevan Skor (1-2)</b>	<b>Relevan Skor (3-4)</b>
<b>Validator</b>	<b>Tidak relevan Skor (1-2)</b>	<b>0</b>	<b>0</b>

2	<b>Relevan Skor (3-4)</b>	0	9
---	-------------------------------	---	---

$$\text{Validasi Isi} = \frac{D}{(A+B+C+D)} = \frac{9}{(0+0+0+9)} = 1$$

Berdasarkan nilai validitas isi yang diperoleh yaitu 1 menunjukkan bahwa terdapat relevansi yang kuat dari kedua validator, dengan koefisien validitas isi lebih besar dari 0,81. Sehingga termasuk kriteria “tinggi” maka dapat dikatakan bahwa relevansi antara kisi-kisi soal meliputi indikator soal, jenis soal, tingkat kognitif, dan butir soal yang dikembangkan adalah valid. Hal demikian menunjukkan bahwa instrumen layak diujicobakan pada tahap selanjutnya meskipun sebelumnya dilakukan beberapa revisi berdasarkan saran kedua validator.

#### *One-to-one*

Tahap ini dilakukan uji coba pertama kepada seorang peserta didik dari kelas X IPA 2 yang memiliki kemampuan tinggi dari temannya yang lain. Peserta didik mengerjakan soal selama 45 menit dalam aplikasi *Google Form* dan dalam pengawasan peneliti. Setelah 45 menit, soal tidak dapat diakses lagi oleh peserta didik. Kemudian peserta didik menuliskan komentar berupa saran dan masukan terkait soal yang dikerjakan.

Komentar yang diberikan berupa tingkat kesukaran soal yang tinggi dan beberapa kalimat yang sulit dipahami, dengan komentar tersebut beberapa butir soal telah direvisi termasuk dari segi tingkat kesukaran, dan kalimat dari beberapa butir soal telah disederhanakan agar mudah dipahami. Hasil revisi ini menghasilkan *Prototype II* yang diujicobakan pada tahap selanjutnya.

#### *Small Group*

Berdasarkan saran dan perbaikan dari tahap *One-to-one* maka dilakukanlah revisi, kemudian hasilnya diujicobakan pada tahap ini. Uji coba diberikan kepada enam peserta didik dari kelas X IPA 2 yang berkemampuan rendah, sedang, dan tinggi. Metode pengerjaan soal sama dengan tahap sebelumnya, peserta didik diberi waktu pengerjaan soal selama 45 menit dalam aplikasi *Google Form* dan dalam pengawasan peneliti. Kemudian keenam peserta didik menuliskan komentar berupa saran dan masukan terkait soal yang mereka kerjakan.

Komentar yang diberikan berupa gambar yang tidak muncul, tingkat kesukaran soal yang tinggi, pertanyaan yang kurang dipahami, dan penggunaan android yang harus

menggunakan kuota untuk mengakses. Berdasarkan komentar tersebut, butir soal kembali direvisi dengan memperhatikan komentar yang diberikan oleh peserta didik. Hasil revisi ini menghasilkan *Prototype III* yang selanjutnya diujicobakan pada tahap akhir, yakni *Field Test*.

#### *Field Test* (Uji coba lapangan)

*Prototype III* merupakan hasil revisi *Small Group* diujicobakan pada tahap ini dengan subjek uji coba peserta didik kelas X IPA 1 yang berjumlah 35 orang. Pengerjaan soal selama 45 menit dalam aplikasi *Google Form* dengan jumlah soal 30 butir. Setelah 45 menit soal tidak dapat diakses lagi oleh peserta didik. Peserta didik memilih satu jawaban yang tepat dari lima pilihan jawaban yang disediakan setiap soal. Hasil pengerjaan soalnya dianalisis untuk mengetahui kualitas soal yang telah dikembangkan.

#### *Analisis Validitas Empiris Instrumen*

Instrumen penilaian kognitif yang telah dikembangkan kemudian dianalisis menggunakan korelasi *product moment pearson*. Hasil analisis butir soal yang telah diujicobakan kepada peserta didik sebagai berikut:

Tabel 3. Analisis Validasi Empirik Butir Soal

Kriteria	Nomor Soal	Jumlah	Persentase
<b>Tidak Valid</b>	1, 6, 9, 10, 11, 12, 13, 14, 15, 17, 18, 20, 21, 22, 30	15	50%
<b>Valid</b>	2, 3, 4, 5, 7, 8, 16, 19, 23, 24, 25, 26, 27, 28, 29	15	50%

Hasil perhitungan  $r_{xy}$  dibandingkan  $r_{tabel}$  *product moment pearson*, dengan taraf signifikan 5% serta derajat kebebasan  $Df = N-2$  yakni 0,317 sebagai  $r_{tabel}$ . Dengan demikian, terdapat butir soal yang memiliki kriteria valid sebanyak 15 atau 50% dan butir soal lainnya memiliki kriteria tidak valid, sebanyak 15 atau 50%.

Hasil ini menunjukkan bahwa jumlah butir soal yang valid, seimbang dengan butir soal yang tidak valid. Sehingga, butir soal yang valid menjadi produk akhir pada penelitian. Instrumen yang valid adalah instrumen yang dapat digunakan untuk mengukur apa yang seharusnya diukur (Sugiyono, 2016). Sehingga, jika instrumen tidak valid sebaiknya dibuang atau diganti. Tindak lanjut dari hasil ini, butir soal yang

dinyatakan valid dapat disimpan ke dalam bank soal dan dapat digunakan kembali untuk tes hasil belajar.

#### *Analisis Reliabilitas Instrumen*

Adapun hasil perhitungan reliabilitas instrumen, sebagai berikut:

Tabel 4. Analisis Reliabilitas Instrumen

<b>Kuder Richardson 20 (KR-20)</b>	<b>Jumlah Butir Soal</b>
0,70	30

Berdasarkan tabel 4, dapat diketahui bahwa koefisien reliabilitas instrumen secara keseluruhan adalah 0,70 dengan kriteria “tinggi (*reliable*)” (Rosidin, 2017). Hal ini berarti bahwa instrumen telah reliabel dengan jumlah butir soal sebanyak 30.

Instrumen yang reliabel belum tentu termasuk kriteria valid, tetapi instrumen yang valid sudah pasti reliabel. Tinggi rendahnya reliabilitas instrumen ditentukan dari hasil perhitungan. Reliabilitas suatu instrumen menunjuk kepada ketepatan dan konsistensi suatu pengukuran yang dilakukan (Rosidin, 2017).

#### *Analisis Daya Pembeda Instrumen*

Berdasarkan uji coba lapangan (*Field Test*) daya pembeda instrumen tes yang dikembangkan diperoleh hasil sebagai berikut:

Tabel 5. Analisis Daya Pembeda Instrumen

<b>Kriteria</b>	<b>Nomor Soal</b>	<b>Jumlah</b>	<b>Persentase</b>
<b>Sangat Jelek</b>	12, 13, 14, 18, 30	5	17%
<b>Jelek</b>	1, 6, 21, 24	4	13%
<b>Cukup</b>	5, 9, 10, 11, 15, 20	6	20%
<b>Baik</b>	3, 4, 17, 22, 25, 26, 27, 28, 29	9	30%
<b>Sangat Baik</b>	2, 7, 8, 16, 19, 23	6	20%

Berdasarkan Tabel 5 hasil analisis daya pembeda, menghasilkan 5 butir soal atau 17% kriteria sangat jelek, 4 butir soal atau 13% kriteria jelek, 6 butir soal atau 20% kriteria cukup, 9 butir soal atau 30% kriteria baik, dan 6 butir soal atau 20% kriteria sangat baik. Butir soal yang memiliki kriteria sangat jelek dan jelek adalah butir soal yang tidak mampu membedakan kemampuan peserta didik. Sedangkan butir soal yang lain telah memiliki kualitas yang baik, dari segi daya pembedanya (Amri & Syahrani,

2018). Butir soal yang memiliki kualitas yang baik dapat disimpan ke dalam bank soal untuk digunakan kembali.

Daya pembeda merupakan analisis yang bertujuan untuk mengukur apakah soal mampu membedakan kemampuan peserta didik, baik yang memiliki kemampuan tinggi maupun memiliki kemampuan rendah, dalam menguasai kompetensi berdasarkan kriteria tertentu. Jika soal dikerjakan oleh peserta didik yang berkemampuan tinggi, maka soal tersebut akan dapat diselesaikan dengan baik. Sebaliknya, jika soal dikerjakan oleh peserta didik yang berkemampuan rendah, maka hasilnya menunjukkan daya pembeda yang rendah juga (Zulfadli, 2017). Apabila suatu butir soal tidak dapat membedakan kemampuan peserta didik berarti ada beberapa kemungkinan, antara lain kunci jawaban butir soal yang tidak tepat, butir soal memiliki dua atau lebih kunci jawaban yang benar, dan pengecoh tidak berfungsi, serta materi yang ditanyakan terlalu sulit (Kadir, 2015).

#### *Analisis Tingkat Kesukaran Instrumen*

Soal yang memiliki kualitas yang baik akan menunjukkan tingkat kesukaran yang proporsional, dalam artian soal tidak terlalu sukar dan juga tidak terlalu mudah (Syamsudduha, 2014). Berdasarkan uji coba lapangan (*Field Test*), tingkat kesukaran instrumen tes yang dikembangkan diperoleh hasil sebagai berikut:

Tabel 6. Analisis Tingkat Kesukaran Instrumen

<b>Kriteria</b>	<b>Nomor Soal</b>	<b>Jumlah</b>	<b>Persentase</b>
<b>Sukar</b>	1, 5, 9, 12, 14, 15, 30	7	23%
<b>Sedang</b>	2, 3, 6, 7, 8, 11, 16, 19, 20, 22, 23, 24, 25, 26, 27, 29	16	54%
<b>Mudah</b>	4, 10, 13, 17, 18, 21, 28	7	23%

Berdasarkan tabel 6 diketahui instrumen yang dikembangkan telah memiliki tingkat kesukaran yang bervariasi dengan kriteria mudah, sedang, dan sukar. Sebanyak 7 butir soal atau 23% yang mudah, 16 butir soal atau 56% yang sedang, dan 7 butir soal atau 23% yang sukar. Soal dalam kualitas yang baik akan menunjukkan tingkat kesukaran yang proporsional. Sehingga secara keseluruhan tingkat kesukaran instrumen yang dikembangkan sudah baik. Soal yang terlalu sukar ataupun terlalu mudah bukan berarti tidak dapat digunakan, akan tetapi penggunaannya dapat disesuaikan dengan

tujuan pemberian soal tersebut (Syamsudduha, 2014). Selain itu, soal dalam kriteria terlalu mudah juga tidak akan membuat peserta didik untuk berusaha dalam memecahkannya. Sebaliknya, soal dalam kriteria terlalu sukar akan membuat peserta didik menjadi mudah putus asa dalam memecahkannya (Taufiq, 2015).

#### *Analisis Efektivitas Pengecoh Instrumen*

Hasil perhitungan efektivitas pengecoh instrumen pada tahap uji coba lapangan (*Field Test*) sebagai berikut:

Tabel 7. Analisis Efektivitas Pengecoh Instrumen

<b>Kriteria</b>	<b>Nomor Soal</b>	<b>Jumlah</b>	<b>Persentase</b>
<b>Tidak Efektif</b>	17, 18, 21, 22, 27, 28	6	20%
<b>Cukup</b>	4, 7, 8, 10, 11, 13, 15, 16, 24, 25	10	33%
<b>Efektif</b>	1, 2, 3, 5, 6, 9, 12, 14, 19, 20, 23, 26, 29, 30	14	47%

Berdasarkan tabel 7 hasil analisis pengecoh yang tersedia pada setiap butir soal yang dikembangkan terdapat 20% yang termasuk kriteria tidak efektif, 33% yang termasuk kriteria cukup, dan 47% lainnya termasuk kriteria efektif. Butir soal yang memiliki kriteria efektif artinya pengecoh bekerja dengan baik, untuk kriteria cukup artinya pengecoh tidak bekerja sepenuhnya, sedangkan untuk kriteria tidak efektif artinya pengecoh tidak bekerja sama sekali. Butir soal yang tidak efektif ini sebaiknya diganti. Hal yang perlu diperhatikan dalam membuat pengecoh pada soal, sebaiknya menggunakan kata-kata yang memiliki kaitan dengan pilihan jawaban yang benar, menggunakan kata-kata yang kedengarannya hampir sama namun tetap memiliki kaitan, dan pilihan jawaban yang mudah dimengerti oleh peserta didik (Arif, 2017).

Pengecoh atau *distractor* bertujuan untuk membuat peserta didik terkecoh. Tujuan lain dari *distractor* pada setiap butir soal adalah membuat peserta didik yang mengikuti evaluasi tertarik untuk memilihnya sebab mereka mengira bahwa *distractor* yang mereka pilih adalah jawaban yang benar (Amri & Syahrani, 2018). Pengecoh dapat dikatakan bekerja dengan baik apabila peserta didik yang memilih pilihan tersebut dan menganggapnya benar.

Tabel 8. Kriteria 30 Butir Soal

No	Validitas Empirik		Daya Pembeda		Tingkat Kesukaran		Efektivitas Pengecoh					Ket.	
	r <sub>htg</sub>	Kriteria	DP	Kriteria	P	Kriteria	IP						
							A	B	C	D	E		Kriteria
1	0,152	Tidak Valid	0,14	Jelek	0,14 3	Sukar	50%	166 %	**	150 %	33%	Efektif	Ditolak
2	0,491	Valid	0,71	Sangat Baik	0,67 9	Sedang	88%	88%	177 %	44%	**	Efektif	Diterima
3	0,509	Valid	0,57	Baik	0,64 3	Sedang	80%	80%	16%	**	80%	Efektif	Diterima
4	0,497	Valid	0,57	Baik	0,78 6	Mudah	133 %	0%	200 %	66%	**	Cukup	Diterima
5	0,349	Valid	0,28	Cukup	0,21 4	Sukar	200 %	90%	109 %	**	54%	Efektif	Diterima
6	0,156	Tidak Valid	0,14	Jelek	0,46 4	Sedang	**	108 %	108 %	81%	108 %	Efektif	Ditolak
7	0,614	Valid	0,85	Sangat Baik	0,46 4	Sedang	243 %	**	81%	0%	81%	Cukup	Diterima
8	0,563	Valid	0,71	Sangat Baik	0,46 4	Sedang	108 %	243 %	0%	54%	**	Cukup	Diterima
9	0,225	Tidak Valid	0,28	Cukup	0,14 3	Sukar	133 %	66%	**	133 %	50%	Efektif	Ditolak
10	0,268	Tidak Valid	0,28	Cukup	0,75 0	Mudah	**	117 %	176 %	117 %	0%	Cukup	Ditolak
11	0,191	Tidak Valid	0,28	Cukup	0,64 3	Sedang	160 %	0%	**	40%	200 %	Cukup	Ditolak
12	-0,01	Tidak Valid	-0,1	Sangat Jelek	0,17 9	Sukar	122 %	87%	70%	122 %	**	Efektif	Ditolak
13	0,158	Tidak Valid	0,00	Sangat Jelek	0,71 4	Mudah	150 %	**	0%	200 %	50%	Cukup	Ditolak
14	0,081	Tidak Valid	0,00	Sangat Jelek	0,10 7	Sukar	48%	112 %	193 %	**	48%	Efektif	Ditolak
15	0,266	Tidak Valid	0,28	Cukup	0,21 4	Sukar	54%	90%	18%	236 %	**	Cukup	Ditolak
16	0,574	Valid	0,71	Sangat Baik	0,53	Sedang	93%	**	93%	180	31%	Cukup	Diterima



					6					%			
17	0,249	Tidak Valid	0,42	Baik	0,78 6	Mudah	4%	**	0%	0,9 %	0%	Tidak Efektif	Ditolak
18	0,000	Tidak Valid	0,00	Sangat Jelek	1,00 0	Mudah	**	0%	0%	0%	0%	Tidak Efektif	Ditolak
19	0,516	Valid	0,71	Sangat Baik	0,50 0	Sedang	200 %	28%	28%	57%	**	Efektif	Diterima
20	0,148	Tidak Valid	0,28	Cukup	0,39 3	Sedang	71%	142 %	47%	**	142 %	Efektif	Ditolak
21	0,041	Tidak Valid	0,14	Jelek	0,85 7	Mudah	200 %	**	0%	0%	200 %	Tidak Efektif	Ditolak
22	0,208	Tidak Valid	0,42	Baik	0,57 1	Sedang	250 %	333 %	333 %	**	83%	Tidak Efektif	Ditolak
23	0,516	Valid	0,71	Sangat Baik	0,50 0	Sedang	200 %	**	142 %	28%	28%	Efektif	Diterima
24	0,322	Valid	0,14	Jelek	0,35 7	Sedang	**	88%	214 %	88%	22%	Cukup	Diterima
25	0,337	Valid	0,42	Baik	0,53 6	Sedang	**	31%	93%	28%	0%	Cukup	Diterima
26	0,41	Valid	0,42	Baik	0,35 7	Sedang	244 %	44%	**	66%	44%	Efektif	Diterima
27	0,446	Valid	0,57	Baik	0,35 7	Sedang	155 %	0%	222 %	22%	**	Tidak Efektif	Diterima
28	0,414	Valid	0,42	Baik	0,78 5	Mudah	**	266 %	0%	66%	0%	Tidak Efektif	Diterima
29	0,444	Valid	0,42	Baik	0,46 4	Sedang	135 %	108 %	108 %	**	54%	Efektif	Diterima
30	-0,11	Tidak Valid	- 0,14	Sangat Jelek	0,07 1	Sukar	370 %	**	50%	100 %	50%	Efektif	Ditolak

Setelah analisis butir yang diujicobakan kepada peserta didik yang terdiri atas 30 butir soal dilakukan pengambilan keputusan terhadap butir soal apakah diterima atau ditolak. Hasil penelitian diperoleh 15 butir soal yang diterima dan 15 butir soal yang ditolak. Soal yang diterima adalah soal yang telah memenuhi kualitas butir soal. Soal yang ditolak tidak memenuhi kriteria butir soal yang berkualitas sehingga perlu direvisi atau dibuang. Soal yang berkualitas dapat membantu guru membedakan peserta didik yang sudah memahami materi yang telah diajarkan dengan peserta didik yang belum memahami materi (Arif, 2017). Oleh karena itu, sangat penting bagi guru untuk mengembangkan instrumen soal yang terstandar untuk memastikan pengukuran yang dilakukan telah benar mengukur yang seharusnya.

Penelitian ini menghasilkan instrumen yang memiliki perbedaan yang khas dibandingkan pengembangan instrumen yang lain, karena instrumen dikembangkan berbasis *Google Form*. Aplikasi ini sebagai alternatif lain pengganti kertas dalam melakukan evaluasi. Sejalan dengan hasil penelitian yang dilakukan Mardiana & Purnanto (2017), *Google Form* dapat dijadikan sebagai alternatif pembuatan evaluasi berdasarkan penilaian guru. Pengaplikasian *Google Form* juga termasuk pemanfaatan teknologi dalam pembelajaran, khususnya kegiatan evaluasi. Kelebihan lain yang dimiliki *Google Form* adalah *template* nya dapat dibagikan kepada orang lain, dengan mengirimkan ke email tujuan. Soal yang dikemas dalam bentuk *Google Form* juga sangat mendukung pembelajaran dalam jaringan pada masa pandemi yang melanda Indonesia sejak Bulan Maret 2020.

## **KESIMPULAN**

Penelitian dan pengembangan instrumen penilaian kognitif berbasis *Google Form* pada sub materi Invertebrata telah dilakukan. Berdasarkan hasil dan pembahasan dari beberapa tahapan pengembangan yang telah dilakukan, dapat disimpulkan bahwa instrumen tes yang dikembangkan tes yang dikembangkan berjumlah 30 butir soal dalam bentuk pilihan ganda telah termasuk dalam kualitas yang baik. Dilihat dari validitas isi sebesar 1,00 dan validitas empirik 50% valid, reliabilitas sebesar 0,70 dalam kriteria tinggi. Sedangkan untuk daya pembeda baik, tingkat kesukaran bervariasi (mudah, sedang dan sukar), dan efektivitas pengecoh yang bekerja.

## DAFTAR PUSTAKA

- Amri & Syahrani. (2018). Pengembangan Tes Sumatif Mata Pelajaran Biologi Semester Ganjil Kelas IX. *Jurnal Biotek (Online)*. Vol. 6, No. 1. <https://doi.org/10.24252/jb.v6i1.5072>
- Arif, W. P. (2017). Pengembangan Tes Kognitif pada Materi Sistem Pernapasan. *Jurnal Biotek*, 5(2), 108-122. <https://doi.org/10.24252/jb.v5i2.4283>
- Arifin, Z. (2013). *Evaluasi Pembelajaran: Prinsip, Teknik, Prosedur*. Cet.V; Bandung: PT Remaja Rosdakarya.
- Batubara, H. H. & Ariani, D. N. (2016). Workshop Penggunaan *Google Form* Sebagai Media Evaluasi Pembelajaran pada Dosen-Dosen Fakultas Studi Islam. *Jurnal Al-Ikhlash (Online)*. Vol. 2, No. 1. <http://dx.doi.org/10.31602/jpai.v2i1.961>
- Djenno, M., Insua, G.M. and Pho, A. (2015), "From paper to pixels: using Google Forms for collaboration and assessment", *Library Hi Tech News*, Vol. 32 No. 4, pp. 9-13. <https://doi.org/10.1108/LHTN-12-2014-0105>
- Hamirul. (2020). *Metode Penelitian dalam Kerangka Patologi Birokrasi*. Muara Bungo: STIA.
- Hidayat. (2017). *Evaluasi Pembelajaran Berbasis Kompetensi*. Makassar: Syahadah.
- Iqbal, M., et al. (2018). Using Google Form for Student Worksheet as Learning Media. *International Journal of Engineering & Technology (Online)*. Vol. 7. No 3.4. <https://www.sciencepubco.com/index.php/ijet/article/view/20125>.
- Kadir, A. (2015). Menyusun dan Menganalisis Tes Hasil Belajar. *Jurnal Al-Ta'dib (Online)*. Vol. 8, No. 2 (<http://ejournal.iainkendari.ac.id/al-tadib/article/view/411>, diakses 19 November 2019).
- Kim, D. dan Park, C. (2012). Gender Differences in the Effectiveness of Google Forms in Class. *Journal of Instructional Pedagogies (Online)*. Vol. 7. <https://www.aabri.com/manuscripts/111033.pdf>, diakses 12 Oktober 2019).
- Mardiana, T. & Purnanto, A. W. (2017). Google Form Sebagai Alternatif Pembuatan Latihan Soal Evaluasi. FKIP Universitas Muhammadiyah Magelang (*Online*). Vol. 2, No. 1 (<http://journal.ummg.ac.id/index.php/urecol/article/view/1582>, diakses 20 September 2019).
- Paidi, Djukri, Yulaikah, S., & Alfindasari, D. (2017). Development of Instrument to Assess Cognitive Process and Product in Biology Senior High School. *International Journal of Environmental & Science Education (Online)*. Vol. 12, No. 8 (<http://www.ijese.net/makale/1935.html>, diakses 12 Oktober 2019).
- Purwati, D. & Nugroho, A. N. P. (2018). Pengembangan Media Evaluasi Pembelajaran Sejarah Berbasis Google Formulir di SMA N 1 Prambanan. *Jurnal Pendidikan*

*dan sejarah (Online)*. Vol 4, No.1 <https://doi.org/10.21831/istoria.v14i1.19398>, diakses 20 September 2019).

- Rahayu, E. E. & Listiyadi, A. (2014). Pengembangan Alat Evaluasi Pembelajaran Berbasis Information and Communication Technology (ICT) pada Materi Mengelola Dokumen Transaksi. *Jurnal Pendidikan Akuntansi (JPAK) (Online)*. Vol.2,No.2. <https://jurnalmahasiswa.unesa.ac.id/index.php/jpak/article/view/9139>, diakses 20 September 2019).
- Retnawati, H. (2016). *Analisis Kuantitatif Instrumen Penelitian (Panduan Peneliti, Mahasiswa, dan Psikometrian)*. Yogyakarta: Parama Publishing.
- Rodriguez, E. (2018). Google Forms in Library Instruction: Creating an Active Learning Space and Communicating with Students. *Scholarship of Teaching and Learning, Innovative Pedagogy (Online)*. Vol. 1, No. 7 ([https://digitalcommons.humboldt.edu/sotl\\_ip/vol1/iss1/7/](https://digitalcommons.humboldt.edu/sotl_ip/vol1/iss1/7/), diakses 12 Oktober 2019).
- Rosidin, U. (2017). *Evaluasi dan Asesmen Pembelajaran*, Edisi Pertama. Cet. I; Yogyakarta: Media Akademi.
- Rosmainar, L. (2017). Analisis Bahan-Bahan Alternatif Pengolahan dalam Pembuatan Kertas. *Jurnal Inkofar (Online)*. Vol. 1, No. 2 (<http://www.politeknikmeta.ac.id/meta/ojs/index.php/inkofar/article/view/18>, diakses 23 September 2019).
- Sugiyono. (2016). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Syamsudduha, S. (2014). *Penilaian Berbasis Kelas Konsep dan Aplikasi*. Yogyakarta: Aynat Publishing.
- Taufiq, A. U. (2015). Pengembangan Tes Kognitif Berbasis Revisi Taksonomi Bloom pada Materi Sistem Reproduksi untuk Siswa SMA. *Jurnal Biotek (Online)*. Vol. 3, No. 2 (<http://journal.uin-alauddin.ac.id/index.php/biotek/issue/view/183>, diakses 22 Juli 2020).
- Tessmer, M. (1993). *Planning and conducting formative evaluations: Improving the quality of education and training*. Psychology Press.
- Yusuf, M. (2015). *Asesmen dan Evaluasi Pendidikan (Pilar Penyedia Informasi dan Kegiatan Pengendalian Mutu Pendidikan)*, Edisi Pertama. Cet. I; Jakarta: Prenadamedia Group.
- Zulfadli. (2017). Pengembangan Tes Kognitif Berbasis Revisi Taksonomi Bloom pada Materi Struktur dan Fungsi Sel untuk Siswa Kelas XI di SMA Kota Tarakan. *Jurnal Pendidikan Biologi Indonesia (Online)*. Vol. 3, No. 2 (<https://media.neliti.com/media/publications/138644>, diakses (Diakses 22 Juli 2020)).