

Pengembangan Instrumen Asesmen *Critical Thinking Skills* pada Materi Sistem Ekskresi

Mahrawi¹, Usman¹, Maulida Nur Avianti^{1*}

¹Jurusan Pendidikan Biologi, FKIP, Universitas Sultan Ageng Tirtayasa, Banten

*E-mail: maulidanuravianti@gmail.com

Abstrak

Penelitian ini bertujuan untuk mengembangkan dan mengetahui kelayakan instrumen asesmen keterampilan berpikir kritis pada materi sistem ekskresi. Penelitian pengembangan ini menggunakan pemodelan Dick and Carey. Teknik pengumpulan data menggunakan wawancara, studi dokumentasi, angket penilaian kelayakan, dan uji coba skala kecil. Hasil penelitian asesmen critical thinking skills pada materi sistem ekskresi ini memenuhi kriteria kelayakan karena memperoleh penilaian sebesar 87,1% (Sangat Layak) dan analisis kuantitatif menggunakan ANATES yaitu, (1) Validitas, soal pilihan ganda valid berjumlah 13 (65%), dan tidak valid 7 butir (35%), pada soal uraian valid berjumlah 7 butir (70%) dan tidak valid 3 butir (30%), (2) Reliabilitas soal pilihan ganda sebesar 0,84 dan soal uraian sebesar 0,86, (3) Tingkat Kesukaran, soal pilihan ganda yang sukar berjumlah 4 butir (20%), sedang 12 butir (60%), dan mudah 4 butir (20%), sedangkan soal uraian sebanyak 10 (100%) tergolong sedang, (4) Daya Pembeda, pilihan ganda sangat baik 17 butir (85%) dan kurang baik 3 butir (15%), sedangkan pada uraian sangat baik 6 butir (60%) dan kurang baik 4 butir (40%), (5) Efektivitas Pengecoh, soal dengan Efektivitas Pengecoh sangat baik 12 butir (15%), baik 11 butir (14%), kurang baik 2 butir (2,5%), jelek 45 butir (56%), dan sangat jelek 10 butir (12,5%).

Kata Kunci: Instrumen asesmen, kemampuan berpikir kritis, Sistem ekskresi.

PENDAHULUAN

Hasil studi PISA (*Programme for International Student Assessment*) yang merupakan program survei komprehensif tingkat internasional yang memberikan penilaian terhadap kemampuan bernalar siswa, pada tahun 2018 menunjukkan pencapaian yang didapat oleh siswa Indonesia dalam PISA sejak beberapa tahun kebelakang amat memprihatinkan (Rahayu, 2016). Pada tahun 2018 Indonesia menempati ranking 71 dari jumlah total 79 negara yang berpartisipasi dalam PISA. Jumlah skor kemampuan sains yang diperoleh oleh Indonesia adalah

396, dimana skor ini jauh dibawah rata-rata skor yang diperoleh Negara yang mengikuti PISA, yaitu sebesar 489. Pencapaian peringkat Indonesia ini selalu konsisten berada pada peringkat bawah, yang menunjukkan bahwa standar kemampuan Indonesia berada di bawah standar masyarakat global. Berbagai upaya telah dilakukan pemerintah Indonesia untuk mengejar ketertinggalan pendidikan di Indonesia.

Pendidikan merupakan hal yang sangat penting pada abad 21 ini untuk dapat menjamin peserta didik agar mempunyai keterampilan belajar dan

berinovasi, keterampilan menggunakan teknologi dan media informasi, serta dapat bekerja, dan bertahan dengan menggunakan keterampilan untuk hidup (*life skills*) (Arifin, 2017). Menghadapi tantangan pendidikan pada abad 21 ini, peserta didik harus dibekali dengan berbagai keterampilan untuk dapat menjawab tuntutan perkembangan zaman. Menurut *framework 21st Century Education* yang dikembangkan oleh *World Economic Forum* (WEF), ada 16 keterampilan penting yang perlu disiapkan dan dimiliki anak-anak agar mereka *survive* dan berhasil masa sekarang. Enam belas keterampilan itu dibagi ke dalam 3 kelompok diantaranya literasi dasar (*foundational literacies*) kompetensi (*competencies*), dan pendidikan karakter (*character qualities*) (Kemendikbud, 2017). Kelompok kompetensi yang disingkat dengan 4C meliputi: *communication, collaboration, critical thinking and problem solving*, serta *creativity and innovation* merupakan kemampuan sesungguhnya yang ingin dituju dengan Kurikulum 2013 sebagai upaya pemerintah Indonesia untuk dapat mengejar ketertinggalan pendidikan dari Negara-negara lainnya.

Kemampuan berpikir kritis (*critical thinking skills*) pada era globalisasi ini sudah menjadi tuntutan yang perlu dimiliki oleh setiap individu, karena kemampuan mengingat saja tidaklah cukup pada proses pembelajaran saat ini. Menurut Arifin (2017) berpikir kritis merupakan berpikir rasional yang memerlukan kemampuan untuk mengevaluasi suatu pernyataan dan mengidentifikasi suatu alasan,

misalnya bukti yang melandasi evaluasi tersebut. Berpikir kritis menuntut upaya keras untuk memeriksa setiap keyakinan atau pengetahuan asumsi berdasarkan bukti pendukungnya dan kesimpulan lanjutan yang diakibatkan. Guru harus mengembangkan keterampilan berpikir pada peserta didik dalam proses pembelajaran maupun penilaian menuju tingkat yang lebih tinggi (Dewi dkk., 2019).

Critical thinking skills yang dimiliki oleh siswa dapat diukur dengan menggunakan instrumen asesmen *critical thinking skills* untuk dapat mengetahui seberapa jauh pembelajaran dapat berhasil dalam meningkatkan *critical thinking skills* yang dimiliki siswa. Asesmen merupakan sebuah penilaian yang bersifat komprehensif untuk dapat mengetahui kinerja peserta didik, atau secara sederhana dapat diartikan sebagai penilaian. Sebagai salah satu usaha untuk dapat meningkatkan kualitas pendidikan, asesmen merupakan hal yang sangat penting dalam pelaksanaan pembelajaran. Pemahaman dan penalaran ilmiahlah yang saat ini sangat ditekankan dalam pembuatan asesmen pembelajaran (Yulia, 2012).

Hasil wawancara yang dilakukan di SMAN 3 Kabupaten Tangerang terhadap guru mata pelajaran Biologi kelas XI yaitu, belum tersedianya instrumen asesmen yang dapat mengukur *critical thinking skills* yang dimiliki oleh siswa. Selain itu asesmen *critical thinking skills* sangat diperlukan untuk melatih kemampuan bernalar siswa dalam menghadapi Asesmen nasional (AN) sebagai pengganti Ujian

Nasional (UN). Materi sistem ekskresi dipilih karena menurut hasil wawancara, sistem ekskresi merupakan materi yang sulit dipahami dan sangat kompleks karena berkaitan juga dengan materi lain seperti pada sistem endokrin dan sistem pencernaan. Menurut Qumillaila et al. (2017) menyatakan bahwa materi sistem ekskresi termasuk ke dalam salah satu materi biologi yang sulit dimengerti oleh siswa, ini disebabkan materi sistem ekskresi pada manusia di dalamnya terdapat serangkaian proses yang dan melibatkan berbagai macam organ dalam tubuh yang sulit dijelaskan. Berdasarkan hal tersebut, maka penelitian ditujukan untuk mengembangkan serta mengetahui kelayakan dari instrumen asesmen *critical thinking skills* pada materi sistem ekskresi, yang diharapkan mampu mengukur kemampuan *critical thinking skills* yang dimiliki oleh siswa.

METODE

Penelitian ini dilaksanakan pada bulan Agustus 2021 yang dilakukan di SMA Negeri 3 Kabupaten Tangerang. Pengembangan instrumen asesmen *critical thinking skills* ini mengacu pada model pengembangan Dick et al. (2001) dengan menggunakan 7 tahapan, yaitu: (1) Mengidentifikasi tujuan umum dengan melakukan analisis kebutuhan yang didasarkan dari informasi guru biologi, (2) Melakukan analisis pembelajaran dengan cara mengumpulkan aktivitas pembelajaran agar tujuan pembelajaran dapat tercapai, (3) Melakukan analisis siswa dan lingkungan dengan merumuskan

kisi-kisi angket sebagai bahan untuk memperoleh informasi mengenai kendala, hambatan, dan kesalahan yang dialami oleh peserta didik dalam mempelajari materi sistem ekskresi, (4) Merumuskan tujuan khusus dengan pemberian tes berupa 25 soal pilihan ganda untuk mengetahui pengetahuan awal peserta didik, (5) Mengembangkan instrumen penilaian dengan merancang dan menyusun pernyataan-pernyataan mengenai penilaian pengembangan asesmen *critical thinking skills* yang meliputi konstruksi, materi, dan bahasa untuk diujikan kepada ahli serta mengembangkan instrumen asesmen *critical thinking skills* yang berisikan 20 butir soal pilihan ganda dan 10 soal uraian pada materi sistem ekskresi, (6) Melakukan evaluasi formatif dengan menguji kelayakannya instrumen asesmen oleh empat orang ahli dan satu orang guru, serta uji coba kelompok kecil, dan (7) Melakukan revisi.

Subjek penelitian yang digunakan mencakup ahli/ pakar, guru biologi, dan 8 orang siswa di SMAN 3 Kabupaten Tangerang.

Teknik pengumpulan data dilakukan dengan wawancara, studi dokumentasi, angket penilaian kelayakan (uji ahli), dan uji coba kelompok kecil. Teknik pengumpulan data yang digunakan yaitu tes dan non tes. Instrumen tes yang dibutuhkan yaitu berupa 25 soal pilihan ganda untuk menguji kemampuan awal siswa serta 20 soal pilihan ganda dan 1-10 soal uraian untuk mengetahui *critical thinking skills* yang dimiliki siswa pada materi sistem ekskresi. Pengumpulan

data non-tes menggunakan angket dan wawancara.

Instrumen asesmen yang sudah divalidasi menghasilkan data kualitatif yang berisi saran dan masukan pada instrumen asesmen untuk perbaikan asesmen. Hasil penilaian dari para ahli dan guru biologi, akan dianalisis untuk mengetahui kelayakan isi instrumen asesmen. Teknik pengolahan data yang diperlukan pada uji coba produk skala kecil pada 8 orang siswa yaitu dengan menggunakan uji validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh butir soal yang nantinya akan dianalisis dengan menggunakan software ANATES.

1. Uji kelayakan ahli

Penilaian kelayakan angket oleh ahli akan dianalisis secara kualitatif menggunakan skala likert meliputi beberapa aspek, yaitu; kualitas materi, konstruksi bentuk soal, bahasa, pedoman hasil analisis jawaban, dan tampilan yang akan dijabarkan dalam angket penilaian kelayakan kemudian akan diubah ke dalam bentuk kuantitatif. Data yang telah didapatkan dari hasil uji ahli kemudian dihitung persentasenya dengan menggunakan rumus Riduwan (2013) sebagai berikut:

$$\text{Persentase kelayakan} = \frac{\text{SkorDiperoleh}}{\text{SkorMaksimal}} \times 100\%$$

Kemudian setelah itu akan dilakukan perhitungan skor rata-rata hasil uji ahli dari beberapa validator dengan rumus (Sudjana, 2009) sebagai berikut:

$$\text{Skor rata-rata} = \frac{\text{Jumlah Skor}}{\text{Jumlah Penilai}} \times 100\%$$

Persentase yang diperoleh kemudian diinterpretasikan dalam kriteria pada Tabel 1 berikut.

Tabel 1. Kriteria Interpretasi Skor

| Persentase | Interpretasi |
|------------|---------------------|
| 81 – 100% | Sangat layak |
| 61 - 80% | Layak |
| 41 – 60% | Cukup layak |
| 21 - 40% | Kurang Layak |
| 1 - 20% | Sangat Kurang Layak |

2. Validitas butir soal

Nilai validitas didapatkan dari bantuan software ANATES. Setiap item soal dan kuesioner dapat diketahui tingkat validitasnya. Kriteria validitas mengacu pada Tabel 2 berikut.

Tabel 2. Kriteria Validitas

| Indeks Korelasi | Interpretasi |
|-----------------|---------------|
| 0.8 – 1,0 | Sangat tinggi |
| 0.6 – 0.79 | Tinggi |
| 0.4 – 0.59 | Cukup tinggi |
| 0.2 – 0.39 | Rendah |
| 0.0 – 0.19 | Sangat rendah |

(Riduwan, 2013: 98)

3. Reliabilitas butir soal

Instrumen bisa dikatakan reliabel apabila memberikan hasil skor yang konsisten pada setiap pengukuran. Nilai yang didapatkan dari ANATES kemudian dikelompokkan ke dalam acuan pada Tabel 3 sebagai berikut.

Tabel 3. Kriteria Realibilitas

| Indeks Korelasi | Interpretasi |
|-----------------|--------------------|
| ≤ 0.20 | Tidak ada korelasi |
| 0.20 – 0.39 | Korelasi rendah |
| 0.40 – 0.69 | Korelasi sedang |
| 0.70 – 0.79 | Korelasi tinggi |

| Indeks Korelasi | Interpretasi |
|-----------------|------------------------|
| 0.80 – 0.99 | Korelasi sangat tinggi |
| 1.0 | Korelasi sempurna |

(Uno & Lamatenggo, 2014: 181)

4. Tingkat kesukaran butir soal

Tingkat kesukaran soal adalah pengukuran seberapa mudah atau sulit suatu soal bagi siswa. Nilai yang didapatkan dari hasil perhitungan menggunakan ANATES, kemudian dikelompokkan ke dalam acuan seperti pada Tabel 4 sebagai berikut.

Tabel 4. Kriteria Indeks Kesukaran Butir Soal

| Indeks Korelasi | Interpretasi |
|-----------------|--------------|
| 0,81 – 1,00 | Sangat mudah |
| 0,61 – 0,80 | Mudah |
| 0,41 – 0,61 | Cukup |
| 0,21 – 0,40 | Sukar |
| 0,00 – 0,20 | Sangat sukar |

(Endrayanto & Harumurti, 2014: 262)

5. Daya pembeda butir soal

Daya pembeda adalah kemampuan soal untuk membedakan antara siswa yang sudah menguasai kompetensi dengan siswa yang belum menguasai kompetensi. Nilai yang didapatkan dari perhitungan menggunakan ANATES, kemudian dikelompokkan ke dalam acuan seperti pada Tabel 5 sebagai berikut.

Tabel 5. Kriteria Daya Pembeda Butir Soal

| Indeks Korelasi | Interpretasi |
|-----------------|------------------------------------|
| 0,40 ke atas | Soal sangat baik |
| 0,30 – 0,39 | Soal baik |
| 0,20 – 0,29 | Soal cukup, soal penentu perbaikan |
| 0,19 ke bawah | Soal kurang baik, soal dibuang |

(Arifin, 2009: 133)

6. Efektivitas pengecoh butir soal

Efektivitas pengecoh butir soal pada asesmen *critical thinking skills* pada materi sistem ekskresi ini menggunakan *software* ANATES V4. Nilai yang didapatkan dari hasil pengujian pengecoh soal menggunakan *software* ANATES V4, kemudian akan diinterpretasikan kedalam tabel kriteria acuan pengecoh soal seperti pada Tabel 6 sebagai berikut.

Tabel 6. Kriteria Indeks Pengecoh (%) Interpretasi Indeks Pengecoh Soal

| Kriteria (%) | Interpretasi |
|------------------------|--------------|
| 76 – 125 | Sangat Baik |
| 51 – 75 atau 126 – 150 | Baik |
| 26 – 50 atau 151 – 175 | Kurang Baik |
| 0 – 25 atau 176 – 200 | Jelek |
| Lebih dari 200 | Sangat Jelek |

(Arifin, 2017: 280)

Efektivitas pengecoh pada setiap butir soal dapat dapat disimpulkan dengan menggunakan skala likert sebagai kriteria acuan seperti pada Tabel 7 berikut.

Tabel 7. Kriteria Penilaian Efektivitas Pengecoh Soal

| Pengecoh yang Berfungsi | Kriteria |
|-------------------------|-------------|
| 4 | Sangat Baik |
| 3 | Baik |
| 2 | Cukup Baik |
| 1 | Kurang Baik |
| 0 | Tidak Baik |

(Sugiyono. 2010: 134-135)

HASIL DAN PEMBAHASAN

Tahap pengembangan produk instrumen asesmen *critical thinking skills* pada materi sistem ekskresi

Pengembangan instrumen asesmen *critical thinking skills* pada materi sistem ekskresi ini, menggunakan pemodelan

Dick and Carey ini dikarenakan tahapannya yang bersifat sistematis dan di dalamnya mencakup siswa, pendidik, materi, dan bahan ajar yang ditujukan agar tujuan pembelajaran dapat tercapai. Hal tersebut sejalan dengan pendapat Thamrin (2014) yang menyatakan bahwa salah satu alasan penggunaan pengembangan dengan menggunakan adaptasi Dick and Carey adalah karena di dalamnya mencakup pembelajar, materi, dan sajian bahan ajar yang digunakan agar tujuan yang sudah ditetapkan dapat tercapai, selain itu pengembangan dengan model ini banyak dipergunakan dalam bidang pendidikan karena bersifat sistematis dan prosedural.

1. Identifikasi Tujuan

Pada tahapan ini dilakukan identifikasi tujuan dengan melakukan analisis kebutuhan (*need assessment*) terkait produk yang akan dikembangkan kepada guru biologi di SMAN 3 Kabupaten Tangerang. Berdasarkan hasil wawancara dapat diketahui bahwa permasalahan yang terjadi yaitu belum tersedianya instrumen asesmen di sekolah yang dapat mengukur kemampuan berpikir kritis yang dimiliki oleh siswa. Hal ini dapat menyebabkan kurangnya kesempatan yang didapat siswa dalam mengembangkan serta mengasah kemampuan berpikir kritis yang dimilikinya.

Tujuan pengembangan instrumen asesmen ini yaitu diharapkan mampu dalam mengukur dan mengasah kemampuan berpikir kritis, sehingga keterampilan siswa dalam pemecahan masalah dapat meningkat, yang nantinya akan membuat siswa dapat terlibat aktif

selama berlangsungnya pembelajaran biologi, khususnya pada materi sistem ekskresi.

2. Analisis Instruksional

Tahapan kedua setelah melakukan identifikasi tujuan adalah melakukan analisis instruksional dengan cara menganalisis kurikulum pada garis besar materi sistem ekskresi, dengan standar kompetensi (SK) yang nantinya akan dikembangkan yaitu setelah mengikuti pembelajaran, siswa dapat menggunakan instrumen asesmen ini sebagai pendekatan salingtemas atau dengan menerapkan prinsip sains untuk memecahkan berbagai masalah yang terjadi di kehidupan sehari-hari.

Selanjutnya yaitu merumuskan kisi-kisi soal yang dijadikan panduan dalam mengembangkan butir soal yang disesuaikan dengan cakupan materi pada konsep sistem ekskresi sebagai pedoman dalam penulisan soal.

Butir soal yang dikembangkan terdiri 25 soal pilihan ganda yang ditujukan untuk mengukur kemampuan awal (*entry behavior*) yang dimiliki siswa. Menurut Herlina et al. (2020) memberikan tes awal pada siswa sebelum pembelajaran dimulai dapat mengidentifikasi perilaku yang dimiliki oleh siswa, yang ditujukan untuk menguji *entry behavior* siswa yang berkaitan dengan tujuan khusus pembelajaran yang perlu dikuasai oleh siswa.

Berdasarkan hasil dari tes kemampuan awal (*entry behaviour*) yang telah dilakukan kepada siswa, dapat diketahui bahwa rata-rata siswa menjawab

salah adalah pada soal dengan jenjang kognitif C4 pada dimensi metakognitif. Menurut Asi (2018) kombinasi antara pengetahuan metakognitif dengan pengetahuan prosedural setara dengan kemampuan berpikir kritis, maka dari itu dapat diketahui bahwa siswa memiliki kemampuan berpikir kritis yang kurang.

3. Analisis Siswa dan Lingkungan

Hasil analisis angket karakteristik siswa yang diberikan pada guru biologi diketahui karakteristik yang dimiliki siswa saat ini adalah siswa kurang dapat berpikir dengan kompleks dan efektif, selain itu saat proses tanya jawab selama berlangsungnya pembelajaran pada umumnya tidak melibatkan siswa secara merata, atau hanya didominasi oleh siswa-siswa tertentu saja. Kurangnya keaktifan siswa selama pembelajaran pada materi sistem ekskresi ini, menyebabkan terciptanya iklim belajar yang kurang kondusif, maka karakteristik siswa yang diharapkan dalam pengembangan instrumen asesmen ini adalah sikap

aktif siswa selama berlangsungnya pembelajaran serta kemampuan siswa dalam berpikir kompleks dalam memecahkan permasalahan.

4. Merumuskan Tujuan Khusus Pembelajaran

Setelah melakukan analisis karakteristik siswa, selanjutnya dilakukan perumusan tujuan khusus pembelajaran yang telah didapatkan dari hasil analisis pembelajaran serta analisis siswa dan lingkungan. Berdasarkan hasil dari tahapan tersebut, dirumuskan tujuan khusus yang perlu dicapai yaitu untuk dapat mengukur dan meningkatkan kemampuan berpikir kritis yang dimiliki siswa pada materi sistem ekskresi menggunakan instrumen asesmen *critical thinking skills*, dengan indikator berpikir kritis Ennis yang dituangkan ke dalam kisi-kisi butir soal (Tabel 8). Indikator yang digunakan berisikan 9 sub kemampuan berpikir kritis yang kemudian dimodifikasi oleh Sakti (2014) menjadi 18 aspek berpikir kritis yang diukur.

Tabel 8. Kisi-kisi butir soal Asesmen *Critical Thinking Skills*

| Indikator Critical Thinking | Sub Indikator Critical Thinking | Aspek yang diukur | Nomor Soal | Jumlah Soal |
|------------------------------------|--|---|------------|-------------|
| 1. Memberikan Penjelasan Sederhana | Memfokuskan pertanyaan | Mengidentifikasi atau memformulasikan suatu masalah | 1, 2, 21 | 3 |
| | | Mengidentifikasi atau memformulasikan kriteria jawaban yang mungkin | 3, 22 | 2 |
| | Menganalisis pernyataan | Mengidentifikasi kesimpulan | 4 | 1 |
| | Bertanya dan menjawab pertanyaan tentang suatu penjelasan. | Apa yang menjadi alasan utama | 5, 6 | 2 |
| | | Mengapa hal tersebut dapat terjadi | 7, 8 | 2 |

| Indikator Critical Thinking | Sub Indikator Critical Thinking | Aspek yang diukur | Nomor Soal | Jumlah Soal |
|---------------------------------|---|---|------------|-------------|
| | | Bagaimana mengaplikasikan kasus tersebut | 9, 23 | 2 |
| 2. Membangun keterampilan dasar | Mempertimbangkan apakah sumber dapat dipercaya/ tidak | Kemampuan memberi alasan | 10 | 1 |
| | | Mempertimbangkan prosedur yang tersedia | 11, 24 | 2 |
| | | Laporan dilakukan oleh pengamat sendiri | 12 | 1 |
| 4. Membuat penjelasan lanjut | Mendeduksi dan mempertimbangkan hasil deduksi Menginduksi dan mempertimbangkan hasil induksi | Kemungkinan dalam penguatan | 25 | 1 |
| | | Menginterpretasikan pernyataan | 13, 14, 26 | 3 |
| | | Menggeneralisasi | 15, 16, 27 | 3 |
| 5. Mengatur strategi dan taktik | Menentukan tindakan | Mendefinisikan masalah | 18 | 1 |
| | | Memilih kriteria yang mungkin sebagai solusi permasalahan | 19 | 1 |
| | | Memutuskan hal-hal yang dilakukan | 20, 30 | 2 |
| Jumlah | | | | 30 |

(Sakti, 2014)

5. Mengembangkan Instrumen Penilaian

Setelah melakukan perumusan tujuan khusus yang telah dituliskan pada kisi-kisi butir soal, langkah selanjutnya yaitu mengembangkan butir-butir soal yang disesuaikan dengan cakupan materi pada konsep sistem ekskresi sebanyak 20 soal pilihan ganda dan 10 soal uraian dengan menggunakan panduan kisi-kisi butir soal. Kemudian dilanjutkan dengan pembuatan lembar penilaian yang ditujukan untuk mengetahui kelayakan dari asesmen *critical thinking skills* yang telah dibuat dengan indikator yang meliputi aspek

materi, konstruksi, dan bahasa yang akan diujikan kepada 5 dosen ahli dan 1 guru biologi.

6. Melakukan Evaluasi Formatif

Instrumen asesmen yang telah dikembangkan diuji kelayakannya oleh lima orang ahli dan satu orang guru biologi SMAN 3 Kabupaten Tangerang. Para ahli nantinya akan diminta untuk dapat memberikan masukan yang menjadi dasar perbaikan produk yang dikembangkan. Aspek yang akan dinilai terdiri dari konstruksi soal, materi,

dan bahasa. Penilaian konstruksi soal berkenaan dengan kesesuaian jenis soal yang disajikan dengan teknik penulisan butir soal. Aspek penilaian dari materi berkenaan dengan konsep/ materi sistem ekskresi apakah sesuai dengan butir soal yang tersedia. Aspek penilaian dalam segi bahasa berkenaan adalah apakah butir soal yang disajikan sudah sesuai dengan kaidah penulisan Bahasa Indonesia yang baik dan benar. Presentase rata-rata semua aspek yang didapatkan dari hasil uji validitas oleh ahli adalah sebesar 87,1%, hasil ini termasuk ke dalam kriteria sangat layak, sehingga dapat dilakukan ke tahap uji coba kelompok kecil untuk uji validitas empiris (Tabel 9).

Tabel 9. Hasil Validasi Ahli Terhadap Asesmen Critical Thinking Skills

| Ahli | Hasil Penilaian (%) | Kriteria |
|--------------------|---------------------|--------------|
| Ahli 1 | 90 | Sangat Layak |
| Ahli 2 | 88.5 | Sangat Layak |
| Ahli 3 | 96.1 | Sangat Layak |
| Ahli 4 | 74.6 | Layak |
| Ahli 5 | 92.3 | Sangat Layak |
| Ahli 6 | 81.5 | Sangat Layak |
| Rata-rata (% ± SD) | 87,1 ± 7,89 | Sangat Layak |

Asesmen *critical thinking skills* yang sudah melalui revisi berdasarkan masukan para ahli kemudian akan diuji cobakan pada kelompok kecil terdiri dari delapan orang siswa kelas XII IPA yang meliputi validitas, reliabilitas, indeks kesukaran, dan daya pembeda butir soal. Menurut Riyani et al. (2017) instrumen yang memiliki validitas empiris adalah instrumen yang telah dilakukan uji berdasarkan pengalaman dengan melakukan uji coba pada siswa sehingga dapat dilihat reliabilitas, tingkat

kesukaran, daya pembeda, dan efektivitas pengecoh soal.

7. Melakukan Revisi

Asesmen yang telah divalidasi kemudian, dilakukan revisi yang didasari oleh masukan-masukan dari para ahli yang diperlukan untuk perbaikan instrumen agar lebih baik lagi. Secara umum hasil penilaian dari para ahli terhadap produk penelitian ini, dapat dikategorikan sangat layak (87,1%), meskipun begitu masih terdapat perbaikan pada butir soal yang meliputi kesesuaian kata kerja operasional dengan jenjang kognitif dalam taksonomi Bloom pada butir soal dan kesesuaian rubrik penilaian pada butir soal.

Revisi pertama yaitu terkait penggunaan kata kerja operasional yang kurang sesuai dengan jenjang kognitif yaitu pada soal nomor 9, 17, dan 28. Pada soal nomor 9 dan 28 redaksi atau gaya bahasa diubah menyesuaikan dengan jenjang kognitif C4 (Menganalisis), sedangkan pada soal nomor 17 soal dirubah total karena tidak sesuai dengan jenjang kognitif C4 (menganalisis). Butir soal nomor 17 sebelum revisi, siswa diminta untuk menjawab penyakit apa yang menyerang paru-paru akibat infeksi kuman. Setelah melalui revisi, pada butir soal siswa disajikan sebuah pernyataan mengenai hormon pada ginjal dan diminta untuk menjawab hormon apa yang dimaksud pada pernyataan tersebut.

Revisi selanjutnya yaitu terkait kesesuaian rubrik penilaian pada butir soal, pada soal nomor 22 dan 26. Rubrik penilaian pada butir soal tersebut, dinilai kurang jelas dan terperinci, sehingga

rubrik penilaian tersebut perlu diperbaiki dan diperjelas kembali agar penskoran setiap jawaban peserta didik nantinya dapat terukur dengan baik.

Tahap Uji Coba Produk Instrumen Asesmen Critical Thinking Skills Pada Materi Sistem Ekskresi

Uji skala terbatas penelitian ini dilakukan pada satu sekolah yaitu SMAN 3 Kabupaten Tangerang dengan jumlah subjek sebanyak 8 orang siswa. Hasil jawaban siswa kemudian akan dianalisis dengan menggunakan ANATES V.4. Tes diagnostik yang dilakukan dapat dijadikan dasar untuk memberikan langkah lebih lanjut yang tepat sesuai dengan kelemahan yang dimiliki oleh siswa (Hadi et al., 2015).

Validitas Butir Soal Instrumen Asesmen Critical Thinking Skills

Menurut Suharso (2012) validitas merupakan pengukuran yang dapat memperlihatkan tingkat ketepatan ukuran dalam meneliti konsep pada suatu instrumen. Instrumen dapat tepat digunakan untuk dijadikan ukuran sebuah konsep jika tinggi tingkat validitasnya, dan jika tingkat validitasnya rendah maka akan menggambarkan bahwa instrumen kurang tepat untuk digunakan (Tjoeng & Indriyani, 2014). Skor pada butir soal akan dikorelasikan dengan skor total pada uji validitas ini. Menurut Riduwan (2013) butir soal dapat dikategorikan ke dalam kategori sangat rendah, rendah, cukup tinggi, tinggi, dan sangat tinggi. Tabel di bawah ini merupakan hasil analisis dari validitas tiap butir soal. Hasil validasi butir soal yang telah disusun disajikan pada Tabel 10 dan Tabel 11.

Tabel 10. Analisis Validitas Butir Soal Pilihan Ganda

| Kriteria | Interpretasi | Persentase (%) | Nomor Soal |
|------------|---------------|----------------|--------------------------|
| 0.8 – 1,0 | Sangat tinggi | 10 | 5, 18 |
| 0.6 – 0.79 | Tinggi | 25 | 7, 8, 9, 12, 14 |
| 0.4 – 0.59 | Cukup tinggi | 35 | 2, 4, 10, 15, 16, 17, 20 |
| 0.2 – 0.39 | Rendah | 15 | 1, 11, 19 |
| 0.0 – 0.19 | Sangat rendah | 15 | 3, 6, 13 |

Tabel 11. Analisis Validitas Butir Soal Uraian

| Kriteria | Interpretasi | Persentase (%) | Nomor Soal |
|------------|---------------|----------------|------------|
| 0.8 – 1,0 | Sangat tinggi | 30 | 1, 7, 10 |
| 0.6 – 0.79 | Tinggi | 30 | 4, 6, 8 |
| 0.4 – 0.59 | Cukup tinggi | 30 | 2, 5, 9 |
| 0.2 – 0.39 | Rendah | 10 | 3 |
| 0.0 – 0.19 | Sangat rendah | 0 | - |

Berdasarkan hasil analisis validitas butir soal pilihan ganda terdapat 14 soal yang memiliki validitas yang tinggi dan 6 soal yang memiliki validitas rendah, sedangkan pada soal uraian terdapat 9 soal yang memiliki tingkat validitas yang tinggi dan 1 soal yang memiliki validitas rendah.

Menurut Sukardi (2011) nilai validitas suatu butir soal yang tidak valid dapat dipengaruhi oleh beberapa faktor menurut yang dibedakan menurut sumbernya, yaitu faktor internal evaluasi, faktor administrasi evaluasi dan skor, dan faktor jawaban dari siswa. Faktor dari internal evaluasi diantaranya, petunjuk tes yang tidak jelas, penggunaan kata-kata yang

sulit dimengerti, penyusunan item soal yang tidak tepat atau sesuai materi, alokasi waktu yang kurang, dan jumlah butir soal yang tidak dapat mewakili sampel materi pembelajaran. Faktor administrasi evaluasi dan skor ini diantaranya, waktu yang terbatas sehingga siswa mengerjakan dengan tergesa-gesa, adanya siswa yang menyontek, penskoran tes yang tidak konsisten, dan adanya joki atau pengganti ketika mengerjakan tes. Faktor jawaban dari siswa yaitu seperti, kondisi psikis, fisik, dan lingkungan siswa ketika mengerjakan tes sedang kurang baik.

Reliabilitas Butir Soal Instrumen Asesmen Critical Thinking Skills

Reliabilitas merupakan suatu koefisien yang digunakan untuk mengetahui sejauh mana alat pengukur atau instrumen dapat dipercaya atau tidaknya, instrumen dapat dikatakan realibilitas jika memiliki hasil yang konsisten jika digunakan untuk mengukur sesuatu yang sama secara berulang-ulang (Khumaedi, 2012). Berdasarkan hasil perhitungan menggunakan aplikasi ANATES V4, keseluruhan nilai yang didapatkan dari butir soal pilihan ganda instrumen asesmen critical thinking skills adalah sebesar 0,84 dan pada soal uraian sebesar 0,86. Menurut Uno & Lamatenggo (2014) pada tabel kriteria reabilitasya, nilai reabilitas sebesar 0,84 dan 0,86 termasuk kedalam kategori sangat tinggi.

Tingkat Kesukaran Butir Soal Instrumen Asesmen Critical Thinking Skills

Tingkat kesukaran ditujukan untuk dapat mengetahui sulit atau mudahnya butir soal atau dapat diartikan juga sebagai

peluang siswa memberikan jawaban yang benar pada suatu tingkat kemampuan (Fatimah & Alfath, 2019). Butir soal yang memiliki angka indeks kesukaran sebesar 0,00 ($P= 0,00$), dapat diartikan bahwa butir soal tersebut dapat dikelompokkan butir soal yang sulit, sebab seluruh siswa tidak ada yang mampu menjawab benar, dan sebaliknya jika butir soal memiliki angka indeks kesukaran butir soal 1, 00 ($P= 1,00$), dapat dikelompokkan ke dalam butir soal yang terlalu mudah, sebab seluruh siswa mampu menjawab butir soal dengan benar (Bagiyono, 2017). Menurut Bagiyono (2017) umumnya butir soal dapat dikatakan baik, jika tidak terlalu sulit dan tidak terlalu mudah. Hasil analisis tingkat kesukaran butir soal disajikan pada Tabel 12 dan Tabel 13.

Tabel 12. Analisis Tingkat Kesukaran Butir Soal Pilihan Ganda

| Kriteria | Interpre-tasi | Persen-tase (%) | Nomor Soal |
|-------------|---------------|-----------------|---|
| 0,61 – 0,80 | Mudah | 20 | 2, 3, 9, 20 |
| 0,41 – 0,60 | Sedang | 60 | 1, 4, 5, 7, 8, 11, 12, 14, 16, 17, 18, 19 |
| 0,21 – 0,40 | Sukar | 20 | 6, 10, 13, 15 |

Tabel 13. Analisis Tingkat Kesukaran Butir Soal Uraian

| Kriteria | Interpre-tasi | Persen-tase (%) | Nomor Soal |
|-------------|---------------|-----------------|-------------------------------|
| 0,61 – 0,80 | Mudah | 0 | - |
| 0,41 – 0,60 | Sedang | 100 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 |
| 0,21 – 0,40 | Sukar | 0 | - |

Menurut Mardapi (2017) Butir soal yang memiliki tingkat kesukaran yang baik adalah soal dengan indeks kesukaran

antara 0,30 sampai dengan 0,70 atau dapat dikategorikan soal dengan tingkat kesukaran sedang. Berdasarkan acuan tersebut, maka soal yang dapat digunakan dan tidak dapat digunakan dapat dilihat pada Tabel 14 untuk soal pilihan ganda dan Tabel 15 untuk soal uraian.

Tabel 14. Soal Pilihan Ganda yang dapat Digunakan Berdasarkan Tingkat Kesukaran

| Kriteria | Interpretasi | Persentase (%) | Nomor Soal |
|-----------------|-----------------------|----------------|--|
| 0,3 – 0,8 | Dapat digunakan | 80 | 1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 14, 16, 17, 18, 19, 20 |
| < 0,3 dan > 0,8 | Tidak dapat digunakan | 20 | 6, 10, 13, 15 |

Tabel 15. Soal Uraian yang dapat Digunakan Berdasarkan Tingkat Kesukaran

| Kriteria | Interpretasi | Persentase (%) | Nomor Soal |
|-----------------|-----------------------|----------------|-------------------------------|
| 0,3 – 0,8 | Dapat digunakan | 100 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 |
| < 0,3 dan > 0,8 | Tidak dapat digunakan | 0 | - |

Daya Pembeda Butir Soal Instrumen Asesmen Critical Thinking Skills

Daya pembeda diperlukan untuk dapat mengetahui kemampuan butir soal untuk dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi yang diujikan (Fatimah & Alfath, 2019). Menurut Dewi dkk. (2019) daya pembeda yang semakin tinggi maka akan

menunjukkan semakin baik pula butir soal tersebut dalam membedakan siswa yang sudah atau belum memahami materi yang diujikan.

Berdasarkan tabel kriteria daya pembeda soal menurut Arifin (2009), soal pilihan ganda yang dikategorikan sangat baik karena memiliki interval 0,40 ke atas adalah soal nomor, 1, 2, 4, 5, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, dan 20 sedangkan untuk kategori kurang baik yang memiliki interval 0,19 ke bawah yaitu, soal nomor 3, 6, dan 11. Butir soal uraian yang dikategorikan sangat baik dengan interval 0,40 ke atas yaitu, soal nomor 1, 5, 6, 7, 8, dan 10 sedangkan butir soal yang dikategorikan cukup baik yaitu, soal nomor 2, 3, 4, dan 9. Hasil analisis daya pembeda butir soal disajikan pada Tabel 16 dan Tabel 17.

Tabel 16. Analisis Daya Pembeda Butir Soal Pilihan Ganda

| Kriteria | Interpretasi | Persentase (%) | Nomor Soal |
|---------------|--------------|----------------|---|
| 0, 40 ke atas | Sangat Baik | 85 | 1, 2, 4, 5, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20 |
| 0,19 ke bawah | Kurang Baik | 15 | 3, 6, 11 |

Tabel 17. Analisis Daya pembeda Butir Soal Uraian

| Kriteria | Interpretasi | Persentase (%) | Nomor Soal |
|---------------|--------------|----------------|-------------------|
| 0, 40 ke atas | Sangat Baik | 60 | 1, 5, 6, 7, 8, 10 |
| 0,20-0,29 | Cukup Baik | 40 | 2, 3, 4, 9 |

Menurut Mansyur et al. (2015) butir

soal yang dapat diterima yaitu butir soal yang memiliki indeks daya pembeda sebesar 0,30 atau lebih, sehingga dapat membedakan antarsiswa yang menguasai materi dan siswa yang tidak menguasai materi, jika butir soal kurang dari indeks tersebut maka perlu dilakukannya revisi. Berdasarkan pendapat tersebut, maka butir soal pilihan ganda dan uraian yang dapat digunakan dapat dilihat pada tabel 18 dan tabel 19.

Tabel 18. Soal Pilihan Ganda yang Dapat Digunakan Berdasarkan Daya Pembeda

| Kriteria | Interpretasi | Persentase (%) | Nomor Soal |
|--------------|-----------------------|----------------|---|
| 0,30 ke atas | Dapat digunakan | 85 | 1, 2, 4, 5, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20 |
| < 30 | Tidak dapat digunakan | 15 | 3, 6, 11 |

Tabel 19. Soal Uraian yang Dapat Digunakan Berdasarkan Daya Pembeda

| Kriteria | Interpretasi | Persentase (%) | Nomor Soal |
|--------------|-----------------------|----------------|-------------------|
| 0,30 ke atas | Dapat digunakan | 60 | 1, 5, 6, 7, 8, 10 |
| < 30 | Tidak dapat digunakan | 40 | 2, 3, 4, 9 |

Efektivitas Pengecoh Butir Soal

Butir soal yang sudah diketahui validitas, reliabilitas, tingkat kesukaran, dan daya pembedanya, untuk selanjutnya perlu juga diketahui efektivitas pengecoh pada butir soalnya. Pengecoh butir soal merupakan alternatif pilihan jawaban

yang telah disediakan oleh pembuat soal yang bertujuan untuk mengecoh peserta didik yang kurang paham dengan materi yang sedang diujikan (Warjul dkk., 2020). Pengecoh yang banyak dipilih oleh siswa, dapat disimpulkan bahwa pengecoh tersebut dapat berfungsi dengan baik. Pengecoh yang dipilih paling sedikit 5% dari jumlah siswa yang mengikuti tes, maka pengecoh tersebut termasuk kedalam kriteria baik (Uno & Koni, 2012). Efektivitas butir pengecoh soal disajikan pada Tabel 20 dan Tabel 21.

Tabel 20. Efektivitas Keseluruhan Pengecoh Butir Soal

Tabel 20. Efektivitas Keseluruhan Pengecoh Butir Soal

| Kategori | Jumlah | Persentase (%) |
|--------------|--------|----------------|
| Sangat Baik | 12 | 15 |
| Baik | 11 | 14 |
| Kurang Baik | 2 | 2,5 |
| Jelek | 45 | 56 |
| Sangat Jelek | 10 | 12,5 |

Tabel 21. Efektivitas Pengecoh Butir Soal

| Kategori | Jumlah | Persentase (%) | Nomor Soal |
|------------------------|--------|----------------|---------------------|
| Sangat Baik | 12 | 15 | |
| (4 Pengecoh berfungsi) | 2 | 10 | 10, 14 |
| Baik | 2 | 2,5 | |
| (3 Pengecoh berfungsi) | 5 | 25 | 7, 15, 16, 18, 19 |
| Kurang Baik | 10 | 12,5 | |
| (2 Pengecoh berfungsi) | 6 | 30 | 4, 6, 8, 11, 12, 17 |
| Jelek | | | |
| (1 Pengecoh berfungsi) | 1 | 5 | 1 |
| Sangat Jelek | | | |
| (0 Pengecoh berfungsi) | 6 | 0 | 2, 3, 5, 9, 13, 20 |

Menurut Warju et al. (2020) pengecoh soal yang tidak dapat bekerja dengan baik sehingga masuk ke dalam kategori buruk ini menunjukkan perbedaan antara pengecoh dengan pilihan jawaban yang lainnya terlalu terlihat mencolok sehingga kurang timbulnya ketrtarikan pada peserta didik untuk memilih pengecoh tersebut, sehingga pengecoh yang masuk ke dalam kategori kurang baik perlu direvisi, sedangkan pengecoh yang masuk ke dalam kategori jelek dan sangat jelek perlu diganti.

Hasil Analisis Critical Thinking Skills Siswa

Hasil analisis skor rata-rata siswa kelas XII SMA Negeri 3 Kabupaten Tangerang dalam menjawab asesmen critical thinking skills pada materi sistem ekskresi adalah

sebesar 49,2% yang mana jika rata-rata tersebut dikategorikan ke dalam tabel kategori persentase *critical thinking skills* (Tabel 22) menurut Setyowati (2011) termasuk ke dalam kategori rendah.

Tabel 22. Kategori Persentase *Critical Thinking Skills*

| Interval Nilai (%) | Kategori |
|--------------------|---------------|
| 81,25<X≤100 | Sangat Tinggi |
| 71,50<X≤81,25 | Tinggi |
| 62,50<X≤71,50 | Sedang |
| 43,75<X≤62,50 | Rendah |
| 0<X≤43,75 | Sangat Rendah |

Rata-rata hasil perhitungan dari masing-masing indikator *critical thinking skills* baik pada butir soal pilihan ganda dan uraian dapat dilihat pada tabel persentase indikator *critical thinking skills*.

Tabel 20. Kategori Persentase *Critical Thinking Skills*

| Indikator Critical Thinking | Sub Indikator Critical Thinking | Aspek yang diukur | Persentase |
|------------------------------------|---|---|------------|
| 1. Memberikan Penjelasan Sederhana | Memfokuskan pertanyaan | Mengidentifikasi atau memformulasikan suatu masalah | 54,4% |
| | | Mengidentifikasi atau memformulasikan kriteria jawaban yang mungkin | 62,5 % |
| | Menganalisis pernyataan Bertanya dan menjawab pertanyaan tentang suatu penjelasan. | Mengidentifikasi kesimpulan | 50% |
| | | Apa yang menjadi alasan utama | 37,5 % |
| | | Mengapa hal tersebut dapat terjadi | 43,8% |
| | | Bagaimana mengaplikasikan kasus tersebut | 60% |
| 2. Membangun keterampilan dasar | Mempertimbangkan apakah sumber dapat dipercaya/ tidak | Kemampuan memberi alasan | 25% |
| | | Mempertimbangkan prosedur yang tersedia | 48,6% |
| | Mengamati dan mempertimbangkan suatu laporan hasil observasi | Laporan dilakukan oleh pengamat sendiri | 75% |
| | | Kemungkinan dalam penguatan | 50% |

| Indikator Critical Thinking | Sub Indikator Critical Thinking | Aspek yang diukur | Persentase |
|---------------------------------|--|---|------------|
| 3. Menyimpulkan | Mendeduksi dan mempertimbangkan hasil deduksi | Menginterpretasikan pernyataan | 33,1% |
| | Menginduksi dan mempertimbangkan hasil induksi | Menggeneralisasi | 31,3% |
| 4. Membuat penjelasan lanjut | Mendefinisikan istilah dan mempertimbangkan definisi | Strategi definisi | 50% |
| | Mengidentifikasi asumsi | Konten | 40% |
| 5. Mengatur strategi dan taktik | Menentukan tindakan | Alasan yang tidak dinyatakan | 45% |
| | | Mendefinisikan masalah | 50% |
| | | Memilih kriteria yang mungkin sebagai solusi permasalahan | 62,5% |
| | | Memutuskan hal-hal yang dilakukan | 66,3% |

Persentase pada aspek memberikan penjelasan sederhana yaitu sebesar 55,9% masuk kedalam kategori rendah. Pada aspek membangun keterampilan dasar sebesar 47,4% dalam kategori rendah. pada aspek menyimpulkan sebesar 40,6% dalam kategori sangat rendah. Pada aspek membuat penjelasan lanjut sebesar 51,6% dalam kategori rendah. Pada aspek mengatur strategi dan taktik sebesar 61,4% dalam kategori rendah.

SIMPULAN

Berdasarkan penelitian ini dapat disimpulkan pengembangan asesmen critical thinking skills pada materi sistem ekskresi yang dikembangkan di SMAN 3 Kabupaten Tangerang ini termasuk ke dalam kategori sangat layak (87,1%). Pada hasil analisis butir soal kedua jenis soal menunjukkan nilai reabilitas yang sangat tinggi yaitu, pada pilihan ganda sebesar 0,84 dan uraian sebesar 0,86. Soal

memenuhi indeks tingkat kesukaran antara 0,3 sampai 0,7, dan daya pembeda sebesar 0,30 atau lebih, maka soal pilihan ganda yang dapat digunakan tanpa perlu dilakukan revisi adalah soal nomor 1, 2, 4, 5, 7, 8, 9, 12, 14, 16, 17, 18, 19, 20 sedangkan soal yang dapat digunakan dengan revisi adalah soal nomor 3 dan 11. Soal uraian yang dapat digunakan tanpa perlu dilakukan revisi adalah soal nomor 1, 2, 5, 6, 7, 8, 10 sedangkan soal yang dapat digunakan dengan revisi adalah soal nomor 2, 3, 4, dan 9. Berdasarkan hasil analisis skor rata-rata siswa dalam menjawab asesmen critical thinking skills adalah sebesar 51,25 yang termasuk ke dalam kategori rendah.

DAFTAR PUSTAKA

- Arifin, Z. (2009). *Evaluasi pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Arifin, Z. (2017). *Mengembangkan Instrumen Pengukur Critical Thinking Skillss Siswa pada Pembelajaran*

- Matematika Abad 21. Jurnal THEOREMS (The Original Research of Mathematics), 1(2), 92-100.
- Bagiyono. (2017). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat 1. *Widyanuklida*, 16(1), 1-12.
- Dewi, S.S., R.M. Hariastuti & A.U. Utami. (2019). Analisis Tingkat Kesukaran Dan Daya Pembeda Soal Olimpiade Matematika (Omi) Tingkat Smp Tahun 2018. *Jurnal Pendidikan Matematika & Matematika*, 3(1), 15-26.
- Dick, W., Carey, L. & Carey, J.O. (2001). *The systematic design of instruction* (5th). New York: Longman.
- Fatimah, L. U. & K. Alfath. (2019). Analisis Kesukaran Soal, Daya Pembeda Dan Fungsi Distraktor. *Jurnal Komunikasi & Pendidikan Islam*, 8(2), 37-64.
- Hadi, S., K.I. Ismara & E. Tanumihardja. (2015). Pengembangan Sistem Tes Diagnostik Kesulitan Belajar Kompetensi Dasar Kejuruan Siswa Smk. *Jurnal Penelitian dan Evaluasi Pendidikan*, 19(2), 168-175.
- Kemendikbud. (2017). Materi Pendukung Literasi Baca Tulis. Gerakan Literasi Nasional. Diakses dari <https://gl.kemdikbud.go.id/glnsite/wpcontent/uploads/2017/10/literasi-baca-tulis.pdf> pada tanggal 5 April 2021.
- Khumaedi, M. (2012). Reliabilitas Instrumen Penelitian Pendidikan. *Jurnal Pendidikan Teknik Mesin*, 12(1), 25-30.
- Mansyur, Rasyid, H., & Suratno. (2015). Asesmen pembelajaran disekolah. Panduan bagi guru dan calon guru. Yogyakarta: Pustaka Pelajar.
- Mardapi, D. (2017). Pengukuran, penilaian, dan evaluasi pendidikan: Edisi 2. Yogyakarta: Parama Publisng.
- Rahayu, S. (2016). Mengembangkan Literasi Sains Anak Indonesia Melalui Pembelajaran Berorientasi Nature Of Science (NOS). Prosiding Seminar Nasional IPA ke-2 yang diselenggarakan Oleh, 17 Maret 2016. Malang: Universitas Negeri Malang.
- Riduwan. 2014. *Metode & Teknik Menyusun Tesis*. Bandung: Alfabeta.
- Setyowati, A. 2011. Implementasi Pendekatan Konflik Kognitif dalam Pembelajaran Fisika untuk Menumbuhkan Kemampuan Berpikir Kritis Siswa Kelas VIII. *Jurnal Pendidikan Fisika Indonesia*, 7(2) : 89-96.
- Tjoeng, S. C. & R. Indriyani. (2014). Pengaruh Perceived Organizational Support Terhadap Corporate Entrepreneurship Pada Perusahaan Keluarga Di Jawa Timur. *AGORA*, 2(1), 1-8.
- Uno, B. H., & Lamatenggo. (2014). *Teori Kinerja dan Pengukurannya*. Jakarta: Bumi Aksara.
- Warjul., S.R. Ariyanto., Soeryanto & R.A. Trisna. (2020). Analisis Kualitas Butir Soal Tipe Hots Pada Kompetensi Sistem Rem Siswa Di Sekolah Menengah Kejuruan. *Jurnal Pendidikan Teknologi dan Kejuruan*, 17(1), 95-104.
- Yulia, R. G. (2012). Pentingnya Asesmen Alternatif Dalam Meningkatkan Kemampuan Berpikir Dan Membaca Ilmiah Siswa Padapembelajaran Biologi. *Jurnal Scientiae Educatia*, 1(1): 1-17.

PROFIL SINGKAT

Maulida Nur Avianti lahir pada tanggal 16 Juli di Kota Pontianak. Ia memulai pendidikan dasar di SD Negeri Curug Kulon pada tahun 2005, SMP Negeri 1 Curug pada tahun 2011, SMA Negeri 3 Kabupaten Tangerang pada tahun 2014 dan masuk ke perguruan tinggi di Jurusan Pendidikan Biologi di Universitas Sultan Ageng Tirtayasa, Banten angkatan 2017.