

## KEEFEKTIFAN PRODUK PENGEMBANGAN INSTRUMEN PENILAIAN KOGNITIF UNTUK MENGUKUR KEMAMPUAN KOGNITIF SISWA (MENURUT TAKSONOMI BLOOM YANG TEREFISI) PADA MATERI PROTISTA

Eny Tarliany<sup>1</sup>, Sajidan<sup>2</sup> Karyanto<sup>3</sup>

<sup>1</sup>Program Studi Magister Pendidikan Sains FKIP Universitas Sebelas Maret  
Surakarta, 57126, Indonesia  
[etarliany@gmail.com](mailto:etarliany@gmail.com)

<sup>2</sup>Program Studi Magister Pendidikan Sains FKIP Universitas Sebelas Maret  
Surakarta, 57126, Indonesia  
[sajidan@fkip.uns.ac.id](mailto:sajidan@fkip.uns.ac.id)

<sup>3</sup>Program Studi Magister Pendidikan Sains FKIP Universitas Sebelas Maret  
Surakarta, 57126, Indonesia  
[karyarina@yahoo.com](mailto:karyarina@yahoo.com)

### Abstrak

Tujuan penelitian yaitu: Menguji keefektifan produk pengembangan instrumen penilaian kognitif untuk mengukur kemampuan kognitif siswa pada materi protista. Pengembangan instrumen penilaian kognitif untuk mengukur kemampuan berpikir tingkat tinggi mengacu pada model *research and development* (R&D) modifikasi dari Borg and Gall meliputi 9 langkah, yaitu: 1) penelitian pendahuluan dan pengumpulan informasi, 2) perencanaan, 3) pengembangan produk awal, 4) uji coba produk awal, 5) revisi produk I, 6) uji coba lapangan, 7) revisi produk II, 8) uji coba lapangan operasional, dan 9) revisi produk akhir. Pada penelitian ini hanya dilakukan sembilan tahapan R&D. Analisis hasil penelitian menggunakan teknik deskriptif dan kualitatif. Hasil penelitian meliputi: 1) produk pengembangan instrumen penilaian kognitif dikembangkan berdasarkan indikator keterampilan proses kognitif menurut Anderson dan Krathwohl 2010, 2) kelayakan isi produk instrumen penilaian kognitif untuk mengukur keterampilan proses kognitif termasuk dalam kategori “sangat baik” dan validitas konstruk “sangat baik”; validitas butir soal dengan kategori interpretasi minimal “cukup”; tingkat kesukaran soal dengan proporsi 11,00% kategori sedang dan 89,00% kategori sulit; daya beda dengan interpretasi minimal “cukup”; kepraktisan soal termasuk kategori “baik”, 3) Perbedaan nilai kemampuan kognitif siswa kelas evaluasi dengan GDL (77,10) dengan kelas *existing learning* dengan DL (70,60) menunjukkan keefektifan produk instrumen penilaian kognitif ini.

**Kata Kunci :** Pengembangan, Instrumen Penilaian Kognitif, Kemampuan Berpikir Tingkat Tinggi, Protista

### Pendahuluan

Pembelajaran abad 21 menuntut siswa memiliki keterampilan-keterampilan yang memadai sesuai dengan dinamika perkembangan globalisasi secara dinamis. Keterampilan-keterampilan yang dimaksud adalah

kemampuan berkomunikasi, berkolaborasi, dan berpikir tingkat tinggi. Ketiga kemampuan tersebut merupakan tuntutan wajib sebagai kompetensi yang harus diperoleh siswa. Secara terintegrasi, komponen tersebut memberi harapan dalam membentuk insan cerdas berkarakter. Kemampuan berpikir tingkat

tinggi adalah penggunaan pikiran secara lebih luas untuk menemukan tantangan baru. Kemampuan berpikir tingkat tinggi ini menghendaki seseorang untuk menerapkan informasi baru atau pengetahuan sebelumnya dan memanipulasi informasi untuk menjangkau kemungkinan jawaban dalam situasi baru (Heong *et al.*, 2011).

Standar Penilaian Pendidikan kurikulum 2013 adalah kriteria mengenai mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik (Permendikbud Nomor 104, Tahun 2014). Penilaian pendidikan sebagai proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik disebut penilaian autentik. Penilaian autentik adalah kegiatan menilai peserta didik yang menekankan pada apa yang seharusnya dinilai, baik proses maupun hasil dengan berbagai instrumen penilaian yang disesuaikan dengan tuntutan kompetensi yang ada di standar kompetensi (SK) dan Kompetensi Dasar (KD), meliputi 3 ranah : kognitif, psikomotor, dan afektif yang dilakukan secara proporsional. Diantara teknik dan instrumen penilaian dalam kurikulum 2013 adalah 1) Penilaian Kompetensi Sikap melalui observasi, penilaian diri, penilaian “teman sejawat” oleh peserta didik dan jurnal. 2) Penilaian Kompetensi Pengetahuan melalui, tes tulis, tes lisan, dan penugasan; 3) Penilaian Kompetensi keterampilan melalui penilaian kinerja. Menurut Asmin (2006), peningkatan mutu pendidikan tidak terlepas dari penerapan penilaian yang dapat secara tepat mengukur hasil akhir dari suatu proses pembelajaran, artinya untuk menilai hasil akhir dalam pembelajaran diperlukan alat ukur yang berkualitas. Salah satu alat ukur adalah tes.

Instrumen tes yang baik dapat meningkatkan kualitas hasil penilaian yaitu dapat mengetahui profil kemampuan peserta didik. Penilaian dengan menggunakan tes tertulis paling

sering digunakan untuk mengetahui kemampuan kognitif siswa. Instrumen pembelajaran yang berorientasi pada keterampilan menjadi penting untuk dikembangkan.

Keterampilan berpikir dapat dikembangkan melalui suatu pengkondisian untuk berpikir. Oleh karena itu, dibutuhkan proses latihan berpikir melalui menjawab soal yang berorientasi pada keterampilan berpikir tingkat tinggi sehingga siswa mampu mengikuti perkembangan ilmu pengetahuan dan teknologi (Yildirim dan Özkahraman, 2011).

Berdasarkan hasil analisis 8 Standar Nasional Pendidikan (SNP) di SMA Negeri 3 Magelang pada standar penilaian menunjukkan gap yang cukup besar antara data ideal (18,06) dengan data real (16,67), adanya gap ini dapat disimpulkan bahwa kurang maksimalnya penyusunan dan pelaksanaan pada tiap-tiap indikator penilaian. Hasil analisis soal ulangan harian menunjukkan bahwa soal-soal yang digunakan kurang melatih berpikir tingkat tinggi, soal katagori C4 hanya 8,3 %, C5 hanya 0 %, C6 hanya 0 % dan soal yang digunakan cenderung pada katagori C1: 46,7 %, C2: 40 % dan C3: 5 %. Berdasarkan fakta di atas, diperlukan suatu bentuk instrumen penilaian kognitif yang dapat mengukur keterampilan berpikir tingkat tinggi siswa.

Berdasarkan latar belakang, maka diperlukan pengembangan instrumen penilaian kognitif yang dapat mengukur keterampilan berpikir tingkat tinggi siswa.

## Metode Penelitian

Model penelitian dan pengembangan (*Research & Development*) yang dilakukan mengacu pada modifikasi dari Borg dan Gall. Langkah pengembangan sebagai berikut : 1) Melakukan penelitian pendahuluan

dan pengumpulan informasi (*research and information collecting*), yaitu mengenali permasalahan yang ada dilapangan, menganalisis proses pembelajaran guru, menganalisis bentuk soal, dan melakukan studi pustaka; 2) Melakukan perencanaan (*planning*), yaitu menentukan kompetensi dasar, merumuskan indikator pembelajaran; 3) Mengembangkan bentuk produk awal (*develop preliminary from of product*), yaitu menyiapkan materi pembelajaran dan membuat kisi-kisi instrumen penilaian; 4) Melakukan *preliminary field testing*, yaitu melakukan validasi bentuk awal dari produk pada ahli dan pengguna dalam skala terbatas; 5) Melakukan revisi terhadap produk utama (*main product revision*), yaitu melakukan perbaikan sesuai dengan saran-saran dari hasil *preliminary field testing*; 6) Melakukan *main field testing*, yaitu memvalidasi produk pengembangan dalam skala lebih luas serta dibandingkan dengan produk apabila memungkinkan; 7) Melakukan revisi terhadap produk operasional (*operational product revision*), yaitu melakukan revisi produk berdasarkan saran-saran dari hasil *main field testing*; 8) Melakukan *operational field testing* yaitu melakukan uji validasi terhadap produk operasional yang dihasilkan; 9) Melakukan revisi terhadap produk akhir (*final product revision*), yaitu melakukan revisi produk seperti disarankan dari hasil *operational field testing*.

Penelitian pengembangan instrumen penilaian ini dilakukan di Sekolah Menengah Atas Negeri 3 Magelang. Sampel yang digunakan adalah siswa kelas X5 dan X6 semester ganjil Tahun Pelajaran 2015/2016. Instrumen pengambilan data yang digunakan meliputi angket untuk analisis kebutuhan, lembar *ceklis* 8 SNP, bank soal, dokumen soal guru, silabus dan RPP, lembar *ceklis* penilaian produk dan kepraktisan. Data analisis kebutuhan dianalisis dengan statistik deskriptif.

Hasil angket dideskripsikan untuk menganalisis kebutuhan pengembangan. Hasil analisis digunakan untuk mempertimbangkan kebutuhan pengembangan. Data penilaian ahli dan guru senior terhadap soal dianalisis dengan teknik deskriptif prosentase (Purwanto,2010). Analisis data dilakukan dengan cara menghitung skor yang dicapai dari keseluruhan aspek yang dinilai kemudian menghitungnya dengan rumus sebagai berikut:

$$N = \frac{K}{Nk} \times 100 \%$$

Keterangan:

N : presentase kelayakan aspek  
 K : skor hasil pengumpulan data  
 Nk : skor maksimal (skor kriteria tertinggi x jumlah aspek x jumlah validator)

Tabel 1 Kriteria Interpretasi Skor Validasi Ahli

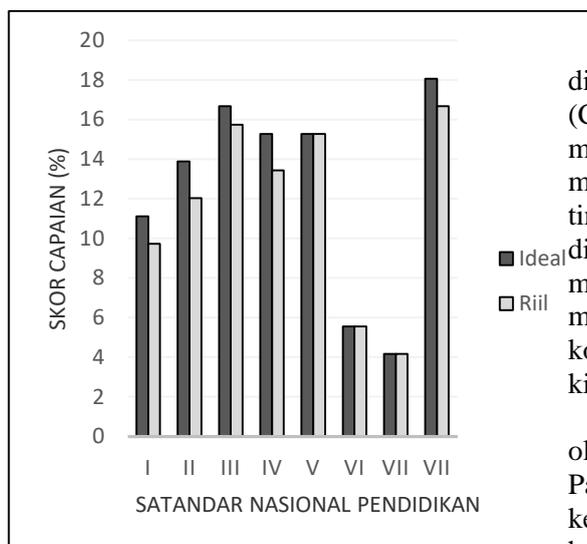
| Interval       | Kriteria      | Konversi |
|----------------|---------------|----------|
| 86% ≤ N < 100% | Sangat Baik   | A        |
| 72% ≤ N < 85%  | Baik          | B        |
| 58% ≤ N < 71%  | Cukup         | C        |
| 44% ≤ N < 57%  | Kurang        | D        |
| N < 44%        | Sangat Kurang | E        |

Validasi dilakukan oleh ahli evaluasi, validasi materi, praktisi pendidikan dan guru bidang studi. Pengumpulan data menggunakan data kualitatif dan data kuantitatif. Data kualitatif diperoleh dari guru bidang studi. Data kuantitatif diperoleh dari pengujian soal meliputi validitas, reliabilitas, daya pembeda dan kepraktisan soal. Teknik yang digunakan dalam penelitian ini meliputi analisis soal-soal tes terkait protista yang ada pada ulangan harian dan bank soal yang digunakan di sekolah. Penelitian deskriptif kualitatif dan kuantitatif dengan *post test only control group desain* (Sugiyono, 2011). Sampel yang digunakan dalam riset ini adalah 48 siswa pada uji lapangan dan 60 siswa pada uji

operasional. Data hasil uji coba dianalisis menggunakan program *Quest* (Rost, 1988).

### Hasil Penelitian dan Pembahasan

Data yang diperoleh dalam penelitian dan pengembangan antara lain data analisis kebutuhan, data validasi ahli dan praktisi, data hasil uji coba lapangan dan data uji operasional. Data analisis kebutuhan meliputi tingkat pemenuhan Standar Nasional Pendidikan (SNP) di SMA Negeri 3 Magelang, analisis bank soal yang digunakan guru. Analisis SNP mengenai 8 Standar Pendidikan Nasional digunakan untuk mengetahui standar yang dicapai oleh SMA Negeri 3 Magelang. Hasil SNP disajikan pada gambar 1



Gambar 1. Perbandingan Prosentase Skor Ideal dan Riil 8 SNP di SMA Negeri 3 Magelang

Hasil analisis 8 SNP menunjukkan adanya GAP pada standar proses 1,85% dan standar penilaian 1,39%. Analisis bank soal yang digunakan guru dilakukan untuk mengetahui prosentase penggunaan tingkatan Taksonomi Bloom. Hasil temuan bank soal guru adalah C1 (46,7%), C2 (40%), C3 (5%), C4 (8,3%),

C5 dan C6 (0%) dan belum mempertimbangkan dimensi pengetahuan yang terdiri dari faktual, konseptual, dan prosedural. Hasil analisis UAN kompetensi dasar protista belum menunjukkan ketuntasan dengan nilai rata-rata 74,75. Hasil analisis soal yang digunakan guru disajikan pada tabel 2

Tabel 2. Prosentase Taksonomi Bloom di SMA Negeri 3 Magelang

| Tingkat Taksonomi Bloom | Jumlah soal | Total soal | Prosentase (%) |
|-------------------------|-------------|------------|----------------|
| C1 (Pengetahuan)        | 28          | 60         | 46,7           |
| C2 (Pemahaman)          | 24          | 60         | 40,0           |
| C3 (Aplikasi)           | 3           | 60         | 5,0            |
| C4 (Analisis)           | 5           | 60         | 8,3            |
| C5 (Evaluasi)           | 0           | 60         | 0              |
| C6 (Mencipta)           | 0           | 60         | 0              |

Sumber : soal guru

Data tersebut menunjukkan bahwa dimensi proses kognitif tingkat tinggi (C4-C6) masih kurang sehingga guru membutuhkan instrumen penilaian untuk mengukur keterampilan berpikir tingkat tinggi siswa. Perencanaan produk dilakukan dengan langkah-langkah: 1) merumuskan kawasan tes, 2) menguraikan materi, 3) menentukan kompetensi yang akan diuji, 4) membuat kisi-kisi soal.

Validasi produk awal dilakukan oleh ahli evaluasi dari dosen ahli evaluasi Pascasarjana UNS untuk memvalidasi kesesuaian soal dengan indikator berpikir tingkat tinggi dan dimensi pengetahuan, ahli materi dosen Biologi Pascasarjana UNS untuk memvalidasi konsep materi sehingga dapat terhindar dari miskonsepsi pada soal yang dikembangkan, dan praktisi oleh guru SMA Negeri 3 Magelang dan SMA Negeri 1 Magelang untuk memvalidasi kelayakan instrumen penilaian kognitif sebelum diuji lapangan.

Hasil penilaian ahli evaluasi untuk aspek konstruk yang meliputi ranah materi dan ranah konstruksi dengan

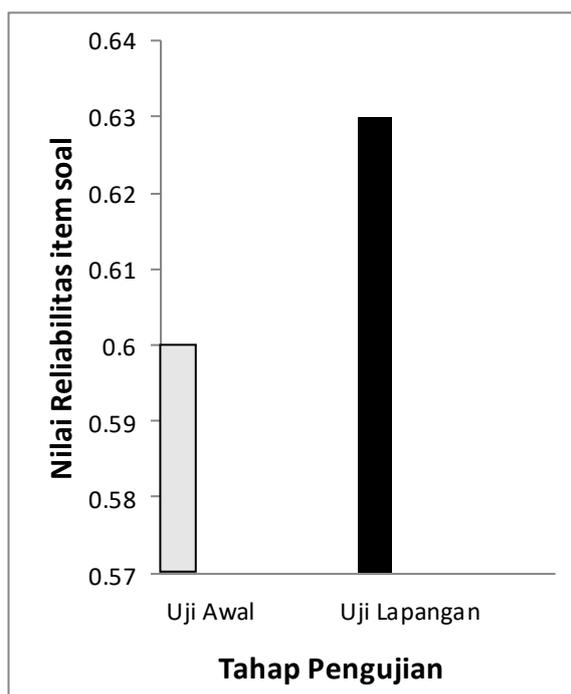
kategori sangat baik (89,99%) dan hasil penilaian indikator konstruk meliputi butir soal sesuai indikator; isi materi yang ditanyakan kontekstual; pokok soal dirumuskan dengan jelas; wacana, gambar atau grafik berfungsi dengan kategori sangat baik (89,99%). Perbaikan dilakukan sesuai saran meliputi gambar yang harus berfungsi, bagan sebaiknya tidak terpotong, dan indikator soal yang baik bersifat paralel.

Hasil penilaian aspek materi oleh ahli materi meliputi kebenaran materi; kedalaman materi; dan keterbacaan soal dengan kategori sangat baik (86,89%) dan hasil penilaian indikator materi meliputi konsep materi soal benar; cakupan materi sesuai tingkatan siswa; istilah yang digunakan jelas; materi soal mudah dipahami; dan materi soal ditulis sistematis, runtut dan alur logika jelas dengan kategori sangat baik (87,14%). Perbaikan dilakukan sesuai saran meliputi gambar yang diperjelas, satuan pada sumbu grafik harus jelas, dan penulisan nama ilmiah makhluk hidup harus sesuai kaidah.

Hasil penilaian praktisi untuk aspek kelayakan instrumen meliputi isi soal; keterbacaan soal, penggunaan bahasa; rubrik penilaian; dan manajemen waktu dengan kategori sangat baik (95,86%) dan hasil penilaian indikator kelayakan instrumen meliputi soal sesuai KD; soal sesuai dengan indikator; soal dapat mengukur kemampuan menganalisis; maksud pertanyaan jelas; perintah mengerjakan soal jelas; istilah yang digunakan jelas; susunan kalimat baik; tidak ada kesalahan tata tulis, ejaan, dan tanda baca; kunci jawaban benar; penskoran objektif; dan waktu siswa cukup untuk mengerjakan soal dengan kategori sangat baik (96,3%). Perbaikan dilakukan sesuai dengan saran praktisi meliputi penulisan ilmiah makhluk hidup harus benar dan menggunakan kalimat yang efektif.

Uji coba lapangan dilakukan pada 1 orang guru dan 48 siswa di SMA

Negeri 3 Magelang. Guru berfungsi untuk menilai kepraktisan instrumen penilaian kognitif. Siswa pada uji coba lapangan diminta untuk mengerjakan soal dan hasilnya dianalisis untuk melihat validitas, reliabilitas, daya beda dan tingkat kesukaran butir soal yang dikembangkan. Menurut Arikunto (2007) sebuah tes yang baik harus memiliki persyaratan kualitas tes, yaitu memiliki validitas, reliabilitas, objektivitas dan praktibilitas. Jumlah item soal adalah 35 dan di uji cobakan sebanyak dua kali. Hasil uji validitas dan reliabilitas instrumen dengan menggunakan program QUEST. Hasil uji awal mendapatkan nilai reliabilitas item soal tinggi (0,6) dan uji coba lapangan (0,63). Hasil reliabilitas disajikan pada gambar 2.



Gambar 2. Hasil Uji Coba Instrumen Penilaian

Hasil uji awal mendapatkan nilai taraf kesukaran kategori mudah 0 %, sedang 11%, dan sukar 89%.

Hasil uji coba lapangan mendapatkan daya pembeda soal dengan interpretasi minimal “cukup” dan

memiliki kepraktisan soal penilaian “baik”. Hasil kepraktisan soal dengan kategori baik meliputi waktu penyusunan tidak lebih dari 1 bulan; dan pelaksanaan tes mudah. Hasil kepraktisan soal dengan kategori baik meliputi biaya penyusunan tes; penyusunan tes dapat dilakukan oleh guru biologi; penilaian tes mudah; mengolah hasil tes mudah; dan waktu pelaksanaan di sekolah cukup. Rata-rata hasil penilaian kepraktisan soal yang dilakukan oleh guru bidang studi termasuk dalam kriteria baik.

Produk pengembangan instrumen penilain yang sudah valid dan reliabel selanjutnya dilakukan uji coba operasional atau keefektifan. Pengujian dilakukan pada dua kelas yang setara. Kelas I (X6) dengan model GDL dan Kelas II (X5) dengan model DL, pada akhir pembelajaran kedua kelas diuji dengan instrumen penilaian kognitif berpikir tingkat tinggi. Hasil perbandingan nilai kemampuan berpikir tingkat tinggi dari kedua kelas disajikan pada tabel 3.

Tabel 3. Perbandingan nilai kemampuan berpikir tingkat tinggi kelas GDL dan kelas DL.

| Kelas | Rata-rata | Perbandingan   |               |
|-------|-----------|----------------|---------------|
|       |           | Nilai Maksimum | Nilai Minimum |
| GDL   | 78,86     | 88,90          | 66,70         |
| DL    | 73,41     | 86,70          | 66,70         |

Perolehan nilai pada kelas GDL memiliki rata-rata 78,86 sedangkan kelas DL 73,41.

Kelas GDL memperoleh nilai diatas KKM. Perbedaan nilai antar dua kelas dipengaruhi oleh beberapa faktor yaitu faktor internal dan faktor eksternal. Faktor internal antara lain kecerdasan, bakat, kematangan dan kesipan siswa dalam menerima instrumen evaluasi. Faktor eksternal antara lain lingkungan sekolah atau lingkungan di dalam kelas (Slameto, 2003).

Instrumen penilaian kognitif yang mampu mengukur keterampilan berpikir

tingkat tinggi memiliki karakteristik, yaitu dikembangkan berdasarkan indikator keterampilan berpikir tinggi dari Anderson dan Krathwohl (2001), meliputi keterampilan menganalisis (membedakan, mengorganisasi, mendekonstruksi), keterampilan mengevaluasi (memeriksa, menilai), keterampilan mencipta (merumuskan, merencanakan) serta dimensi pengetahuan meliputi faktual, konseptual, dan prosedural (Permendikbud 104, 2014).

Produk dari Instrumen penilaian kognitif untuk mengukur keterampilan berpikir tingkat tinggi mampu memberikan stimulus atau rangsangan kepada siswa, tanggapan terhadap adanya rangsangan adalah respon siswa dalam menjawab dengan memilih pilihan jawaban yang telah disediakan.

Hasil penilaian produk oleh ahli evaluasi, ahli materi dan praktisi diperoleh bahwa produk pengembangan instrumen penilaian kognitif termasuk pada kategori baik dan layak digunakan. Rincian hasil penilaian dapat dijelaskan sebagai berikut:

Produk pengembangan instrumen penilaian dinilai oleh ahli evaluasi dengan skor 89,99 %. Berdasarkan hasil penilaian dapat disimpulkan bahwa instrumen penilaian kognitif dapat diimplementasikan dalam pembelajaran di kelas dengan kategori sangat baik.

Produk pengembangan instrumen penilaian dinilai oleh ahli materi dengan skor 87,14 %. Berdasarkan hasil penilaian dapat disimpulkan bahwa instrumen penilaian kognitif dapat diimplementasikan dalam pembelajaran di kelas dengan kategori sangat baik.

Produk pengembangan instrumen penilaian kognitif dinilai oleh praktisi 95,86 %. Berdasarkan hasil penilaian dapat disimpulkan bahwa instrumen penilaian kognitif dapat diimplementasikan dalam pembelajaran di kelas dengan kategori sangat baik.

Tahap mengujikan soal di lapangan dengan menggunakan satu guru pengguna dan 48 siswa di SMA Negeri 3 Magelang bertujuan untuk menguji validitas, reliabilitas, daya pembeda, tingkat kesukaran soal yang dikembangkan. Hasil analisis soal menunjukkan bahwa soal yang dikembangkan termasuk kategori cukup. Butir soal memiliki validitas dengan interpretasi minimal cukup, reliabilitas cukup, tingkat kesukaran soal dengan proporsi 11,00 % sedang dan 89,00 % sukar. Hasil perbandingan kualitas instrumen pada tingkat kesukaran kategori mudah menurut (Kurniasari., *et all*, 2013) adalah 67,7%; (Nofiana., *et all*, 2013) adalah 15%. Tingkat kesukaran kategori sedang menurut (Kurniasari., *et all*, 2013) adalah 22,5%; (Nofiana., *et all*, 2013) adalah 80 %. Tingkat kesukaran kategori sukar menurut (Kurniasari., *et all*, 2013) adalah 10%; (Nofiana., *et all*, 2013) adalah 5%; Hasil ini berbeda dengan proporsi tingkat kesukaran soal yang biasa digunakan dengan proporsi 25% soal mudah, 50% soal sedang, dan 25% soal sukar. Perbedaan ini dikarenakan pada penelitian ini soal-soal yang dikembangkan menggunakan indikator keterampilan berpikir tingkat tinggi. Daya pembeda soal dengan interpretasi minimal cukup, serta tingkat kepraktisan soal baik, yaitu 78,57 %. Hasil ini sesuai dengan Arikunto (2007) yang menyatakan bahwa untuk kriteria instrumen yang baik adalah memiliki validitas, reliabilitas, dan daya pembeda minimal cukup

Keefektifan produk instrumen penilaian kognitif dilihat dari nilai tes siswa yaitu hasil belajar kognitif. Keterampilan berpikir tingkat tinggi bukanlah sebuah hasil belajar yang instan dan langsung didapat, melainkan butuh proses. Sejalan dengan penelitian Woolf (2005) menyatakan bahwa dibutuhkan sebuah proses dan latihan yang tidak singkat untuk dapat mengubah ketrampilan berpikir seseorang. Belajar

untuk mengembangkan ketrampilan berpikir akan berhasil apabila banyak dilakukan latihan atau ulangan (Sagala, 2011). Menurut penelitian Treagust (2012) menyatakan bahwa banyak soal yang dapat digunakan untuk menguji pemahaman siswa serta mengukur kemampuan kognitif pada level yang lebih tinggi. Penelitian lain dari Downing dan Haladyna (1989) menyatakan soal pilihan ganda dapat meningkatkan penilaian pembelajaran dan ketrampilan berpikir yang lebih mendalam.

Temuan di lapangan menunjukkan bahwa hasil pengembangan instrumen penilaian kognitif untuk mengukur keterampilan berpikir tingkat tinggi memiliki keunggulan dan kelemahan, yaitu sebagai berikut :

1. Keunggulan instrumen penilaian kognitif:
  - 1) Jumlah materi yang dapat ditanyakan relatif lebih banyak dibandingkan dengan materi yang dicakup soal bentuk uraian
  - 2) Dapat mengukur jenjang kemampuan berpikir tingkat tinggi (analisis, evaluasi, dan mencipta)
  - 3) Penskoran mudah, cepat, dan objektif
  - 4) Dapat digunakan sebagai alat diagnosis pemahaman materi siswa
  - 5) Dapat digunakan untuk mengetahui efektifitas pembelajaran yang dilakukan oleh guru
2. Kelemahan instrumen penilaian kognitif:
  - 1) Penyusunan soal yang baik memerlukan waktu yang relatif lama dibandingkan dengan bentuk soal yang lainnya
  - 2) Kurang dapat digunakan untuk mengukur kemampuan verbal
  - 3) Siswa belum terbiasa menggunakan soal dalam bentuk pilihan ganda bertingkat

## Kesimpulan dan Rekomendasi

Kesimpulan dari penelitian pengembangan instrumen penilaian:

1. Kelayakan produk instrumen penilaian kognitif untuk mengukur kognitif siswa dijamin melalui validitas isi yang dinilai sangat baik oleh ahli materi, validitas konstruk yang dinilai sangat baik oleh ahli instrumen evaluasi, validitas butir soal dengan interpretasi minimal cukup, memiliki tingkat kesukaran soal dengan proporsi 11,00 % sedang dan 89,00 % sulit, memiliki daya beda soal dengan interpretasi minimal cukup serta memiliki tingkat kepraktisan soal yang dinilai baik.
2. Perbedaan nilai kemampuan berpikir tingkat tinggi siswa kelas GDL dengan kelas DL menunjukkan keefektifan produk instrumen penilaian kognitif ini

Rekomendasi dari penelitian pengembangan instrumen ini antara lain :

1. Siswa sebelumnya telah mendapat materi protista.
2. Guru sebaiknya membelajarkan materi pada KD tersebut menggunakan model pembelajaran yang memberdayakan keterampilan berpikir tingkat tinggi.
3. Alokasi waktu yang diberikan untuk mengerjakan 35 soal bentuk pilihan ganda dan pilihan ganda bertingkat dua tidak kurang dari 60 menit.
4. Instrumen penilaian kognitif yang terkait dengan model pembelajaran yang memberdayakan keterampilan berpikir tingkat tinggi masih perlu dikembangkan dan diteliti lebih lanjut.

## Daftar Pustaka

Anderson, L. W., & Krathwohl, D. R., 2001. *A Taxonomy for Learning, Teaching, and Assesing*. New York: Longman. New York:

- Longman.
- Arikunto, S, 2007. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Asmin, 2006. Pengaruh Ragam Bentuk Tes Obyektif dan Gaya Berfikir terhadap Fungsi Informasi Tes: Penelitian Quasi Eksperimental dengan Analisis Item Response Theory di SMU DKI Jakarta. *Jurnal Pendidikan Dan Kebudayaan*, 062 (12), 633–655.
- Borg, W. R., & Gall, M. D., 1983. *Educational Research: An Introduction* (4th ed.). White Plains: Logman Inc.
- Heong, Y. M., Othman, W. B., Yunos, J. Bin, Kiong, T. T., Hassan, R. Bin, Mohaffyza, M., & Mohamad, B., 2011. The Level of Marzano Higher Order Thinking Skills among Technical Education Students. *International Journal of Social Science and Humanity*, 1(2), 121–125.
- Kurniasari, I. R., Sunarmi, & Nugraningsih., 2014. *Pengembangan Instrumen Penilaian Kognitif Materi Gerak Tumbuhan dan Hama Penyakit Tumbuhan Kelas VIII Semester Genap Tahun Ajaran 2013-2014*.
- Nofiana, M., Sajidan, & Karyanto, P. 2013. *Pengembangan Instrumen Evaluasi Two-Tier Multiple Choice Question untuk Mengukur Keterampilan Berpikir Tingkat Tinggi*.
- Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia No 104 Tahun 2014: *Penilaian Hasil Belajar oleh Pendidik pada Pendidikan Dasar dan*

*Pendidikan Menengah*. Jakarta.

Peraturan Menteri Pendidikan dan Kebudayaan No 66 Tahun 2016: Standar Penilaian Pendidikan, (Standar Penilaian Pendidikan).  
<https://doi.org/10.1017/CBO9781107415324.004>

Purwanto, N, 2010. *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya.

Sagala, S, 2011. *Konsep dan Makna Pembelajaran*. Bandung: Alfabeta.

Sudrajat, A, 2013. Pendekatan Saintifik dalam Proses Pembelajaran.

Treagust, D, 2012. Diagnostic assessment in science as a means to improving teaching, learning and retention. *Australian Conference on Science*, (1998), 1–9. Retrieved from <http://openjournals.library.usyd.edu.au/ind>

Woolf, B. P, 2005. Distributed Interactive Learning Environments New pedagogy: Inquiry learning.

Yildirim, B., & Özkahraman, Ş., 2011. Critical Thinking in Nursing Process and Education. *International Journal of Humanities and Social Science*, 1(13), 257–262.