



PENGEMBANGAN INSTRUMEN TES TIPE PILIHAN GANDA DALAM EVALUASI HASIL BELAJAR SISWA PADA KONSEP SEL TINGKAT SEKOLAH MENENGAH ATAS

Maulida^{1*}, Syarifah Farisi Hamama²

^{1,2}Program Studi Pendidikan Biologi, Universitas Abulyatama, Aceh Besar, Indonesia.

*Email korespondensi : maulida_biologi@abulyatama.ac.id¹

Diterima Januari 2021; Disetujui Januari 2021; Dipublikasi 31 Januari 2021

Abstract: *This study aims to obtain a multiple choice type test instrument that conforms to the standard standards for cell material in high school biology subjects. This research was conducted from February to April 2014. The method used was a formative research type development method with a multiple choice test of 100 test items. The research subjects used were 43 students. Tests that were compiled and validated by a team of experts were given to the research subjects, then the data obtained were analyzed using the formula for difficulty level, different tests (different power), the effectiveness of the trickster, validity, and reliability. The results showed that the number of valid test items was 70 test items and had a high category of reliability. From the results of this study, it can be concluded that the test items developed have met the good test criteria, this is based on the categories in which all the requirements have been fulfilled.*

Keywords : *Development of test instruments, multiple choice, cells.*

Abstrak: Penelitian ini bertujuan untuk memperoleh instrumen tes tipe pilihan ganda yang sesuai standar baku untuk materi sel pada pelajaran biologi tingkat sekolah menengah atas. Penelitian ini dilakukan pada bulan Februari sampai dengan April 2014. Metode yang digunakan adalah metode pengembangan tipe *formative reseach* dengan bentuk tes pilihan ganda sebanyak 100 item tes. Subjek penelitian yang dipakai adalah sebanyak 43 siswa. Tes yang disusun dan telah divalidasi oleh tim ahli diberikan kepada subjek penelitian, kemudian data yang didapatkan dianalisis dengan menggunakan rumus tingkat kesukaran, ujibeda (dayapembeda), efektivitas pengecoh, validitas, dan reliabilitas. Hasil penelitian menunjukkan bahwa jumlah item tes yang telah valid sebanyak 70 item tes dan memiliki kategori tinggi pada reliabilitasnya. Dari hasil penelitian ini, dapat disimpulkan bahwa item tes yang dikembangkan telah memenuhi kriteria tes yang baik, hal ini didasarkan pada kategori dalam tahapan yang telah dilaksanakan telah terpenuhi semua syarat-syaratnya.

Kata kunci : *Pengembangan instrumen tes, pilihan ganda, sel.*

Evaluasi adalah proses akhir dalam sebuah proses pembelajaran, dimana menjadi titik poin untuk melihat keberhasilan proses pembelajaran yang telah dilakukan. Namun untuk melakukan evaluasi bukanlah proses yang mudah, banyak tahapan guna mendapatkan sebuah evaluasi yang baik.

Valid atau tidak nya sebuah evaluasi akan sangat berimbas bagi proses pembelajaran.

Bahan ujian atau soal yang bermutu dapat membantu pendidik meningkatkan pembelajaran dan memberikan informasi dengan tepat tentang peserta didik mana yang belum atau sudah

mencapai kompetensi.

Evaluasi proses pembelajaran dapat dilakukan dengan berbagai bentuk, baik dengan cara melakukan tes dan cara non tes. Salah satu bentuk instrumen tes yang dapat digunakan mengukur tingkat kognitif adalah pilihan ganda. Pilihan ganda merupakan jenis evaluasi yang sering dipakai oleh guru dalam pembelajaran, baik dalam ujian harian maupun ujian akhir. Akurasi data kemampuan siswa atau data kesulitan siswa dalam belajar sangat tergantung kepada akurasi alat evaluasi dan proses evaluasi. Oleh karena itu, alat evaluasi harus disusun secermat mungkin, agar secara konsisten mampu mengukur apa yang semestinya diukur. Selain itu, alat evaluasi harus diujicoba, dan bila perlu harus diujicoba beberapa kali, agar persyaratan validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan persyaratan alat evaluasi lainnya dapat dipenuhi dengan baik.

Pilihan ganda sering dianggap sebuah bentuk instrument tes yang sangat mudah dalam hal penyusunan serta dalam proses pengerjaan yang dilakukan peserta didik. Sering dinggap tidak kompeten dalam melakukan tugasnya dalam evaluasi suatu proses pembelajaran.

Tujuan penelitian ini adalah untuk mengetahui pengembangan instrumen tes tipe pilihan ganda pada materi sel pada materi sel ditingkat sekolah menengah atas.

KAJIAN PUSTAKA

Menurut Permana (2011) “Secara empirik sebuah tes yang digunakan untuk evaluasi pembelajaran dapat dikatakan baik dan benar apabila memiliki ciri-ciri sebagai berikut:

1. Tingkat kesukaran, yaitu angka yang

menunjukkan proporsi siswa yang menjawab betul suatu soal. Makin besar tingkat kesukaran berarti soal itu makin mudah demikian juga sebaliknya yaitu makin rendah tingkat kesukaran berarti soal itu makin sukar.

2. Validitas, yaitu ketepatan mengukur yang dimiliki oleh sebuah item dalam mengukur apa yang seharusnya diukur lewat butir item tersebut.
3. Daya pembeda, yaitu kemampuan suatu butir item tes hasil belajar untuk dapat membedakan antara *testee* yang berkemampuan tinggi dan *testee* yang berkemampuan rendah, demikian juga sehingga sebagian *testee* yang memiliki kemampuan yang tinggi untuk menjawab butir item tersebut lebih banyak menjawab butir item tersebut lebih banyak yang menjawab betul, sementara *testee* yang berkemampuan rendah untuk menjawab butir item tersebut secara besar tidak dapat menjawab item dengan betul.
4. Reliabilitas, yaitu suatu instrumen cukup dapat dipercaya untuk digunakan sebagai alat pengumpul data karena instrumen tersebut sudah baik atau dapat diandalkan”.

Langkah-langkah pengembangan tes meliputi (1) menentukan tujuan penilaian, (2) menentukan kompetensi yang diujikan (3) menentukan materi penting pendukung kompetensi (urgensi, kontinuitas, relevansi, keterpakaian), (4) menentukan jenis tes yang tepat (tertulis, lisan, perbuatan), (5) menyusun kisi-kisi, butir soal, dan pedoman penskoran, (6) melakukan telaah butir soal. Penilaian non tes dilakukan melalui pengamatan dengan langkah-langkah (1) menentukan tujuan penilaian, (2) menentukan

kompetensi yang diujikan, (3) menentukan aspek yang diukur, (4) menyusun tabel pengamatan dan pedoman penskorannya, (5) melakukan penelaahan (Anonymous, 2008).

METODE PENELITIAN

Metode penelitian yang digunakan adalah metode penelitian pengembangan atau *development research* tipe *formative research* (Lewy, 2009). Penelitian pengembangan tipe *formative reseach* melalui beberapa tahapan yaitu:

Preliminary

Pada tahap ini adalah menentukan tempat dan subjek penelitian dengan cara menghubungi Kepala Sekolah dan guru mata pelajaran biologi di sekolah yang akan dijadikan lokasi penelitian serta mengadakan persiapan-persiapan lainnya, seperti mengatur jadwal penelitian dan prosedur kerjasama dengan guru kelas yang akan dijadikan tempat penelitian.

Tahap *Self Evaluation*

a. Analisis

Tahapan analisis ini merupakan langkah awal penelitian pengembangan. Peneliti dalam hal ini menganalisis siswa, analisis materi, kurikulum dan literatur., dalam hal ini peneliti akan mulai menganalisis kurikulum yang berlaku sekarang, materi sel karena sel adalah materi yang dipilih peneliti, dan menganalisis buku teks yang akan dipakai untuk memudahkan peneliti dalam pembuatan instrument nantinya.

b. Desain

Langkah selanjutnya adalah penyusunan instrument yang berupa soal-soal tentang materi sel. Pada tahap ini, peneliti mendesain soal-soal untuk mengukur kemampuan berpikir pada konsep

sel. Desain produk ini sebagai *prototype*. Masing-masing *prototype* fokus pada tiga karakteristik yaitu: konten, konstruks dan bahasa. Ketiga karakteristik ini divalidasi oleh pakar dan teman sejawat. Cara ini dikenal dengan teknik triangulasi.

Prototyping (validasi, evaluasi dan revisi)

Pada tahap ini produk yang telah dirancang kemudian dievaluasi. Dalam tahap evaluasi ini produk akan diujicobakan. Ada 3 kelompok uji coba ini :

a. *Expert Review dan One-to-one*

Hasil desain pada *prototype* pertama yang dikembangkan atas dasar *self evaluation* diberikan pada pakar (*expert review*) dan seorang siswa (*one-to-one*) secara paralel. Dari hasil keduanya dijadikan bahan revisi.

Pada tahap uji coba pakar disini atau biasanya disebut uji validitas, produk yang telah didesain akan dicermati, dinilai dan dievaluasi oleh pakar. Pakar-pakar akan menelaah konten, konstruks dan bahasa dari masing-masing *prototype*. Pada tahap ini, tanggapan dan saran dari para validator tentang desain yang telah dibuat, saran-saran validator ditulis pada lembar validasi sebagai bahan merevisi dan menyatakan bahwa soal-soal untuk mengukur kemampuan berfikir tersebut telah valid. Dari hasil saran para pakar, maka peneliti memperbaiki kembali soal-soal tentang materi sel yang telah dirancang.

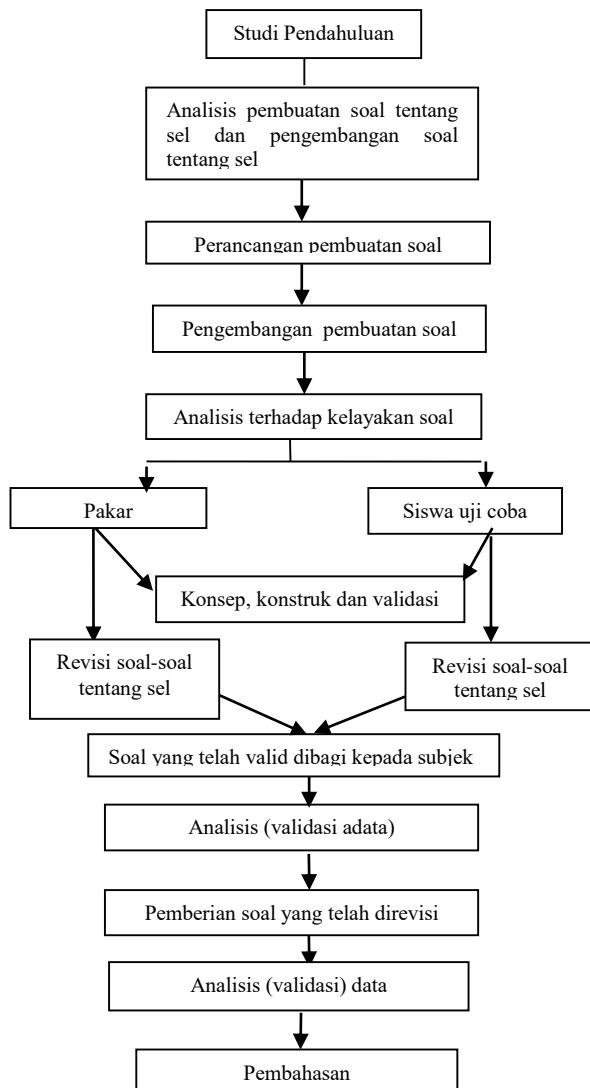
b. *Small Group (kelompok kecil)*

Hasil revisi dari *expert* dan kesulitan yang dialami siswa saat ujicoba pada prototipe pertama dijadikan dasar untuk revisi desain prototype pertama yang dinamakan prototipe ke dua. Kemudian hasilnya diujicobakan pada *small group* (5 orang siswa sebaya non subjek penelitian). Pada

tahap ini diminta 5 orang siswa untuk menyelesaikan soal tentang sel yang telah didesain ulang oleh peneliti. Berdasarkan hasil tes dan komentar siswa soal-soal tersebut kembali direvisi dan dirancang ulang.

3. Field Test (Uji lapangan)

Saran-saran serta hasil uji coba pada prototipe ke dua dijadikan dasar untuk merevisi desain *prototype* kedua. Hasil revisi diujicobakan ke subjek penelitian dalam hal ini sebagai *field test*



Gambar 1. Langkah-langkah penyusunan item tes
 Populasi dalam penelitian ini adalah seluruh siswa kelas XI Semester I pada 2 SMA. Sampel penelitian diambil sebanyak 43 siswa yang

dijadikan sebagai sampel dalam penelitian ini.

Prosedur dan Analisis Data

Menurut Ridhani (2007), data yang telah terkumpul (hasil jawaban siswa) kemudian diolah dengan cara:

Pertama, menentukan jawaban benar dan salah secara manual terhadap semua butir soal dengan berpedoman pada kunci jawaban yang dibuat oleh guru. Semua data jawaban siswa akan ditabulasikan dalam bentuk satu nol, jawaban benar diberi skor satu dan jawaban yang salah diberi skor nol.

Kedua, untuk mendapatkan item tes yang valid maka langkah-langkah yang dilakukan adalah menguji tingkat kesukaran, daya pembeda, efektifitas pengecoh, validitas, reliabilitas. Tahapan yang harus dilakukan adalah sebagai berikut:

1. Untuk mencari nilai indeks kesukaran, maka diperlukan rumus:

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

Keterangan:

P = Indeks kesukaran

B = Jumlah siswa yang menjawab soal dengan benar

JS = Jumlah seluruh siswa peserta tes

Dengan kategori:

Tabel 1. Indeks Kesukaran

Indeks Kesukaran	Keterangan
0,00 – 0,30	Soal sukar
0,31 – 0,70	Soal sedang
0,71 – 1,00	Soal mudah

(Arikunto, 2008)

2. Untuk mencari nilai daya beda, maka diperlukan rumus:

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

Keterangan:

ID = Indek Daya Beda

BA = Jumlah kelompok Atas yang menjawab benar

BB = Jumlah kelompok bawah yang menjawab benar

J_A = Jumlah peserta kelompok atas
 J_B = Jumlah peserta kelompok bawah

Dengan kriteria pengambilan keputusan:

Tabel 2. Daya Beda

Daya beda	Keterangan
Negatif	Tidak baik, dibuang
0,71 – 1,00	Baik sekali
0,41 – 0,70	Baik
0,21 – 0,40	Cukup
0,00 – 0,20	Jelek

3. Untuk mencari nilai validitas, dilakukan dengan menggunakan rumus validitas butir:

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

Keterangan:

N = jumlah siswa

X = skor jawaban siswa per item

Y = total skor jawaban siswa yang benar

(Arikunto, 2008)

Kriteria pengambilan keputusan:

Tabel 3. Koefisien Validasi

Koefisien Validitas	Keterangan
0,00 – 0,20	Sangat Rendah
0,21 – 0,40	Rendah
0,41 – 0,60	Sedang
0,61 – 0,80	Tinggi
0,81 – 1,00	Sangat Tinggi

HASIL DAN PEMBAHASAN

Setelah data dari beberapa kali uji coba didapatkan maka terlihat banyak hal yang masih harus direvisi. Item tes yang disusun mewakili ranah antara C1 sampai C5 meliputi ranah mengingat (C1) sebanyak 5 item tes, memahami (C2) sebanyak 21 item tes, mengaplikasi (C3) sebanyak 15 item tes, menganalisis (C4) sebanyak 29 item tes dan evaluasi (C5) sebanyak 2 item tes.

Dalam mendapatkan item tes yang telah valid akan melewati beberapa tahapan yang meliputi:

Validasi Ahli

Hasil validasi yang telah peneliti lakukan

dengan tim ahli (dalam hal ini dosen yang berkompeten di bidangnya), didapatkan banyak saran yang diberikan oleh ahli. Kebanyakan dari saran yang diberikan oleh tim ahli adalah pada bagian redaksi soal atau pada bagian isi soal, banyak redaksi soal yang belum baik, hal ini dikarenakan soal yang baik adalah soal yang memiliki pengantar terlebih dahulu sebelum *point* yang akan ditanyakan, kemudian perbaikan dari segi isi ataupun konstruk dilakukan sesuai dengan apa yang telah disarankan oleh tim ahli, dari *option* pengecoh yang diberikan banyak hal yang harus direvisi, untuk menghindari tidak berfungsinya pengecoh dengan semestinya, dikarenakan pengecoh yang baik adalah pengecoh yang mampu mengecoh siswa yang mengerti akan konsep tersebut dan juga siswa yang tidak begitu memahami konsep tersebut.

Terdapat beberapa soal terletak kesalahan konsep di dalamnya dan harus direvisi kembali agar tidak terjadinya kesalahan konsep yang akan ditanyakan dalam soal. Terdapat satu soal yang masuk kategori salah konsep dan akhirnya soal tersebut diganti dan juga diperbaiki secara keseluruhan, agar tidak menimbulkan miskonsepsi bagi siswa yang membacanya. Terdapat beberapa soal yang ternyata berulang atau soal yang sama namun terletak pada nomor yang berbeda, soal tersebut dibuang dan selanjutnya diganti dengan soal yang baru. Setelah soal tersebut peneliti revisi maka soal tersebut di cek kembali oleh tim ahli sampai soal tersebut dinyatakan sudah layak untuk di ujicoba baru soal tersebut peneliti pakai untuk diberikan kepada subjek ataupun responden.

Reliabilitas

Pada tahapan uji pertama didapatkan 52 item

tes bisa dipakai dan 48 item masih harus direvisi. Nilai reliabilitas KR-21 berada dalam kategori tinggi (Maulida, 2018).

Indeks Kesukaran

Hasil uji pertama terlihat terdapat 52 item tes dengan kategori indeks kesukaran “mudah” hal ini berarti sekitar 52% item tes yang diberikan berada dalam kategori mudah, 40 item tes dengan kategori indeks kesukaran “sedang” atau sekitar 40% item tes berada dalam kategori sedang dan 8 soal berada dalam kategori indeks kesukaran “sukar” atau sekitar 8% soal berada dalam indeks kesukaran sukar (Tabel 1).

Tabel 4. Rekapitulasi Indeks Kesukaran Item Tes

Kategori	Jumlah Item Tes
sukar	8
sedang	40
mudah	52

Indeks kesukaran selalu berbanding terbalik dengan taraf kesukaran dari suatu item tes, sebagai contoh terdapat pada item tes nomor 1, terlihat indeks kesukaran menunjukkan angka 0,84, suatu angka yang tergolong besar namun ditilik dari segi taraf kesukaran, item tes ini tergolong mudah karena banyak subjek penelitian yang dapat menjawab item tes tersebut, demikian juga sebaliknya semakin rendah angka yang ditunjukkan oleh indeks kesukaran maka semakin tinggi pula taraf kesukaran item tes tersebut hal ini dikarenakan banyak subjek penelitian yang tidak mampu untuk menjawab soal tersebut, terlihat pada soal nomor 24, 26, 51, 54, 75, 77, 89, dan soal nomor 99. Hal ini didukung dengan pernyataan Suryabrata (2005), “kelemahan utama indeks kesukaran item tes adalah antara indeks kesukaran item tes berbanding terbalik dengan

taraf kesukaran item tes, artinya makin tinggi indeks kesukaran suatu item tes maka makin rendahlah taraf kesukaran item tes tersebut”.

Daya Beda

Daya beda yang didapatkan dari proses analisis terlihat menunjukkan 19 item tes berada dalam kategori “baik” atau sekitar 19 % item tes tergolong memiliki daya beda yang baik, 33% item tes berada dalam kategori daya beda “cukup”, dan 48% item tes berada dalam kategori daya beda “jelek” Tabel 2.

Tabel 5. Rekapitulasi Daya Beda Item Tes

Kategori	Jumlah Item Tes
Baik	19
Cukup	33
jelek	48

Tujuan dilakukannya daya beda adalah untuk melihat kemampuan yang dimiliki oleh siswa karena disini siswa akan digolongkan kedalam 2 golongan yakni golongan atas dan golongan bawah berdasarkan rekapitulasi jawaban yang telah diberikan oleh siswa.

Efektivitas Pengecoh

Distractor yang terdapat pada item tes terlihat efektif hal ini dikarenakan siswa memilih *option* jawaban tersebut melebihi dari 5%, dan ini artinya bahwa option pengecoh benar-benar menjalankan fungsinya dengan baik yaitu mengecoh siswa yang memiliki konsep kuat terhadap materi dan juga mampu membingungkan bagi siswa yang masih belum terlalu memahami konsep tersebut.

Validitas

Analisis validasi soal juga dilakukan, dalam hal ini terlihat bahwa terdapat 12 item tes atau 12% item tes yang memiliki kategori indeks validasi “tinggi”, 29 item tes atau 29% item tes berada

dalam kategori indeks validasi “sedang”, 24 item tes atau 24% item tes berada dalam kategori indeks validasi “rendah”, 35 item tes atau 35% item tes berada dalam kategori indeks validasi “sangat rendah”.

Tabel 6. Rekapitulasi Validasi Item Tes

Kategori	Jumlah Item Tes
Tinggi	12
Sedang	29
Rendah	24
Sangat Rendah	35

Untuk mendapatkan nilai validitas hal ini berdasarkan kemampuan siswa dalam menjawab per item tes kemudian akan dikalkulasikan dengan jumlah total benar jawaban siswa tersebut baru akan diperoleh nilai validitas untuk 1 nomor item tes demikian juga seterusnya. Kemampuan pribadi siswa sangat mempengaruhi nilai dari validitas item tes, selain itu juga jumlah total benar dari seluruh item tes akan mempengaruhi nilai dari validitas. Jika terdapat banyak siswa yang dapat menjawab item tes dengan benar maka nilai validitas item tes akan berada pada kategori yang tinggi, begitu juga jika sedikit siswa yang mampu menjawab item tes tersebut maka nilai validitas akan turun demikian juga dengan kategori validitasnya.

KESIMPULAN

Dari hasil penelitian ini, dapat disimpulkan bahwa bahwa item tes yang dikembangkan telah memenuhi kriteria tes yang baik, hal ini didasarkan pada kategori dalam tahapan yang telah dilaksanakan telah terpenuhi semua syarat-syaratnya. item tes yang telah valid sebanyak 70 item tes dan memiliki kategori tinggi pada reliabilitasnya.

DAFTAR PUSTAKA

- Arikunto, S. (2008). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara
- Anonymous. (2008). *Panduan Penulisan Butir Soal*. Jakarta: Departemen Pendidikan Nasional.
- Lewy. (2009). Pengembangan Soal Untuk Mengukur Kemampuan Berpikir Tingkat Tinggi Pokok Bahasan Barisan Dan Deret Bilangan Di Kelas IX Akselerasi SMP Xaverius Maria Palembang. *Jurnal Pendidikan Matematika*, Volume 3. No.2, Desember 2009 hal: 14-28.
- Mardapi, D. (2012). *Pengukuran Penilaian*. Yogyakarta: Nuha Medika.
- Maulida. (2015). Analisis Indeks Kesukaran dalam pengembangan item tes pada konsep sel. *Jurnal Edubio Tropika*, 3 (1).
- Maulida., Lubis, S. P. W. (2018). Analisis Tingkat Reliabilitas pada Item Tes Tipe Pilihan Ganda. *Jurnal Dedikasi pendidikan, Vol. 2, No 2 Juli 2018: 128-133*.
- Permana, B. (2011). *Cara Menganalisis Butir Soal Pilihan Ganda dan Uraian*. Online.
- Sudjana, N. (2006). *Penilaian Hasil Belajar Proses Belajar Mengajar*. Bandung: PT.Remaja Rosdakarya.
- Suryabrata, S. (2005). *Pengembangan Alat Ukur Psikologis*. Yogyakarta: C.V Andi Offset
- Susetyo, B. (2011). *Menyusun Tes Hasil Belajar*. Bandung: C.V Cakra.
- Tandiling, E. (2012). Pengembangan instrumen untuk mengukur kemampuan Komunikasi matematik, pemahaman matematik, dan *selfregulated Learning* siswa dalam pembelajaran matematika Di sekolah menengah atas. *Jurnal Penelitian Pendidikan* Vol. 13 No. 1 April 2012.

▪ *How to cite this paper :*

Mualida & Hamama, S.F. (2021).
Pengembangan Instrumen Tes Tipe
Pilihan Ganda Dalam Evaluasi Hasil
Belajar Siswa Pada Konsep Sel Tingkat
Sekolah Menengah Atas. Jurnal
Dedikasi Pendidikan, 5(1), 171–178.