

## **PENGEMBANGAN ASESMEN AUTENTIK PADA MATERI SISTEM PERNAPASAN**

Yati Nurhayati<sup>1</sup>, Bibin Rubini<sup>2</sup>, Indarini Dwi Pursitasari<sup>3</sup>

<sup>1</sup>*Program Studi Pendidikan IPA, Sekolah Pascasarjana Universitas Pakuan*

\*E-mail: *memeyanaku@gmail.com*

**Abstrak:** Penelitian ini bertujuan untuk menghasilkan asesmen autentik yang berupa instrumen penilaian kognitif, afektif dan psikomotor yang valid juga reliabel. Metode penelitian yang digunakan dalam penelitian ini adalah metode pengembangan atau Research and Development (R&D) dengan desain ADDIE (Analysis, Design, Development, Implementation, Evaluation). Instrumen yang digunakan dalam penelitian ini terdiri dari lembar validitas ahli yang digunakan untuk memvalidasi asesmen autentik oleh ahli sebanyak 6 dosen dan guru IPA sebanyak 20 orang, lembar angket untuk mengetahui respon siswa terhadap asesmen autentik. Untuk mengetahui kualitas instrumen yang dikembangkan uji butir soal meliputi tingkat validitas, reliabilitas, daya pembeda dan tingkat kesukaran. Hasil penelitian menunjukkan bahwa 1) hasil validasi pada komponen isi, kontruksi, kebahasaan dan praktikalitas ahli oleh dosen dan guru IPA valid atau layak digunakan. 2) hasil uji butir soal kognitif terdapat 15 soal pilihan ganda dan 8 soal uraian valid dengan reliabilitas tinggi sehingga layak digunakan. Untuk instrumen penilaian sikap terdapat 28 pernyataan valid dengan reliabilitas tinggi yang dapat dipergunakan. Sedangkan pada instrumen penilaian kinerja 7 aspek penilaian yang valid dengan reliabilitas cukup tinggi dapat digunakan dan pada instrumen penilaian produk 5 aspek penilaian valid dengan tingkat reliabilitas rendah, dapat digunakan 3) penerapan instrumen penilaian kognitif, afektif dan psikomotor yang valid dan reliabel dapat mengukur tingkat pencapaian ketuntasan belajar 4) tanggapan siswa terhadap asesmen autentik sebesar 91,02 % (sangat baik).

**Kata Kunci:** Asesmen Autentik, Aspek Kognitif, Afektif, dan Psikomotor

### **PENDAHULUAN**

Pendidikan nasional berfungsi mengembangkan kemampuan dan membentuk watak serta peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa, Untuk mewujudkannya pemerintah menetapkan delapan standar minimal dalam pelaksanaan pendidikan, salah satunya yaitu standar penilaian pendidikan. Standar penilaian pendidikan adalah kriteria mengenai mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik. Penilaian pendidikan diperlukan sebagai proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik yang mencakup penilaian autentik. Penilaian autentik merupakan penilaian yang

dilakukan secara komprehensif untuk menilai mulai dari masukan (*input*), proses, dan keluaran (*output*) pembelajaran. Menurut Permendikbud, No. 104 Tahun 2014 Pasal 1 Ayat 2 penilaian autentik adalah bentuk penilaian yang menghendaki peserta didik menampilkan sikap, menggunakan pengetahuan dan keterampilan yang diperoleh dari pembelajaran dalam melakukan tugas pada situasi yang sesungguhnya. Ini berarti, guru dalam menilai hasil belajar peserta didik harus mencakup penilaian sikap, pengetahuan, dan keterampilan.

Penilaian dapat dilakukan dengan berbagai teknik/cara, seperti penilaian unjuk kerja (*performance*), penilaian sikap, penilaian tertulis (*paper and pencil test*), penilaian proyek, penilaian produk, penilaian melalui kumpulan hasil kerja/karya peserta didik (*portfolio*) dan penilaian diri. Menurut Brown (2004) asesmen merupakan suatu proses yang berkelanjutan untuk mengumpulkan data pembelajaran. Bentuk dan cara penilaian dalam banyak hal memberikan pengaruh penting bagi proses pembelajaran, seperti cara guru memberi pembelajaran dan cara siswa harus belajar, keduanya menentukan capaian kompetensi (Nurgiyantoro, 2008). Oleh karena itu penilaian hendaknya memfokuskan pada proses kegiatan pembelajaran.

Atac (2012) mendefinisikan asesmen autentik sebagai bentuk asesmen dengan menggunakan aktivitas dan tugas yang mencerminkan tujuan pembelajaran, kurikulum sesuai dengan *real life situation*. Asesmen ini menekankan pada evaluasi yang bermakna dalam pembelajaran, yang menggunakan bermacam-macam bentuk asesmen yang menggambarkan pelajaran, kemampuan, motivasi dan sikap siswa yang relevan dengan aktivitas kelas. Menurut Basit dan Wardana (2016) produk perangkat asesmen autentik yang telah dikembangkan dapat digunakan secara maksimal jika guru dan siswa mempelajari terlebih dahulu agar memudahkan dalam proses pembelajaran. Produk ini juga dapat dimanfaatkan untuk kalangan lebih luas dengan penyesuaian terutama penyesuaian pada karakteristik siswa sebagai pengguna. Johnson (2002) memaparkan bahwa asesmen autentik memberikan keleluasan pada siswa untuk dapat menunjukkan apa yang telah mereka pelajari selama pembelajaran. Metode penilaian autentik diyakini sangat tepat untuk diimplementasikan dalam proses pembelajaran karena metode ini memiliki banyak dampak positif, terutama kepada siswa. Meskipun teknik penilaian autentik ini lebih fokus pada peningkatan kualitas hasil belajar siswa daripada prestasi belajar siswa (Ermawati & Hidayat, 2017). Hal ini sesuai dengan pendapat yang disampaikan oleh Gulikers (2004) bahwa asesmen autentik merupakan asesmen yang mengharuskan siswa untuk menggunakan beberapa kompetensi, atau kombinasi pengetahuan, keterampilan, dan sikap, yang perlu mereka praktikkan dalam *professional life*.

Berdasarkan data hasil studi lapangan di temukan bahwa pelaksanaan penilaian autentik di salah satu sekolah menengah di kabupaten sukabumi belum sepenuhnya dilaksanakan oleh guru. Hal ini disebabkan guru kurang memahami tentang penilaian autentik, guru berpendapat terlalu rumit serta membuang waktu juga energi, instrumen yang dibutuhkan dalam penilaian autentik masih kurang memadai dan guru masih mengalami kesulitan dalam membuat penilaian autentik. Selama ini guru-guru cenderung melaksanakan penilaian yang hanya mengukur pengetahuan saja, sedangkan dalam menilai afektif dan psikomotor kurang diperhatikan. Dengan demikian, penentuan nilai akhir atas ketercapaian hasil belajar siswa kurang holistik. Menurut Boadu, *et al* (2014) lokakarya, pelatihan dan peningkatan kapasitas secara reguler tentang penilaian kelas harus diatur untuk para guru, hal ini bertujuan memperbaiki penggunaan penilaian mereka sebagai alat untuk meningkatkan pengajaran dan pembelajaran. Meskipun Rukmana & Mundilarto, (2016) menyatakan dalam hasil penelitiannya bahwa tidak terdapat hubungan antara pelatihan yang telah diterima pendidik dan keterlaksanaan penilaian autentik. Dengan demikian keterlaksanaan penilaian autentik dapat dipengaruhi oleh bagaimana seorang guru berkeinginan untuk menciptakan suatu penilaian yang dapat melihat kemampuan siswa dalam berbagai aspek secara nyata. Penilaian autentik adalah nilai sebenarnya dengan pengukuran yang bermakna secara signifikan atas hasil belajar peserta didik untuk ranah sikap, keterampilan, dan pengetahuan. Dengan demikian perlunya penilaian autentik karena penilaian ini dirasa lebih adil dalam mengukur kemampuan siswa karena tidak hanya di lihat dari satu aspek kemampuan saja melainkan menyeluruh dari berbagai aspek baik kognitif, afektif maupun psikomotor.

Hasil observasi yang dilakukan peneliti terhadap belajar siswa, ditemukan bahwa kebanyakan siswa memiliki kesulitan dalam mempelajari materi IPA, karena luasnya cakupan materi dan konsep IPA yang harus dipelajari siswa. Pembelajaran akan lebih bermakna jika guru berperan sebagai pembimbing dan siswa diberi kesempatan untuk terlibat secara aktif dalam menemukan konsep dari fenomena yang ada di lingkungan (Rizal, 2014). Hasil proses belajar siswa akan menentukan alat evaluasi yang telah direncanakan oleh guru dan menampilkan hasil yang optimal untuk penilaian yang diberikan guru. Menurut Liandari, dkk. (2017) tuntutan kompetensi dalam kurikulum 2013 meliputi pengetahuan, keterampilan, dan sikap akan tercapai melalui pembelajaran dengan pendekatan keterampilan proses yang memungkinkan siswa dapat menumbuhkan sikap ilmiah untuk mengembangkan keterampilan-keterampilan yang mendasar, sehingga dalam proses pembelajaran siswa dapat memahami konsep yang dipelajarinya. Apabila proses penilaian yang biasa dilakukan guru selama ini hanya mampu menggambarkan aspek penguasaan konsep peserta didik, akibatnya sasaran belajar IPA belum dapat dicapai secara menyeluruh. Untuk itu perlu

diupayakan suatu teknik penilaian yang mampu mengungkap aspek produk maupun proses.

Perlunya sebuah inovasi penilaian yang dapat menilai aspek produk dan proses siswa dalam mengembangkan kemampuan berpikir, meningkatkan keingintahuan, serta menuntun siswa dalam memecahkan berbagai masalah pada materi yang diajarkan. **Pengembangan Asesmen Autentik** diharapkan dapat menciptakan suatu penilaian yang dapat meningkatkan keaktifan siswa dalam pembelajaran IPA sehingga dapat mencapai kompetensi yang diinginkan.

## **METODE**

Metode yang digunakan dalam penelitian ini adalah metode penelitian dan pengembangan atau *Research and Development* (R&D). Penelitian dan pengembangan atau *Research and Development* (R&D) adalah metode penelitian yang digunakan untuk menghasilkan produk tertentu dan menguji keefektifan produk tersebut. Adapun desain pengembangan yang digunakan adalah ADDIE (*Analysis, Design, Development, Implementation, Evaluation*).

Pada analisis (*analysis*), dilakukan analisis kebutuhan asesmen autentik dan pemahaman materi pada sistem pernapasan, studi kurikulum yang digunakan di SMP mengenai kompetensi dasar, indikator, materi pembelajaran, strategi pembelajaran dari konsep materi sistem pernapasan, ttudi literatur dengan melakukan analisis beberapa jurnal yang berkaitan dengan rencana penelitian untuk mengumpulkan temuan riset dan informasi lain yang bersangkutan dengan pengembangan produk yang direncanakan.

Pada tahap desain (*design*), yaitu : menentukan kompetensi inti, kompetensi dasar dan indikator yang ingin dicapai. menentukan teknik-teknik penilaian yang digunakan, membuat kisi-kisi dari indikator pengetahuan dan keterampilan yang di rancang, pembuatan rancangan instrumen *asesmen autentik*.

Pada tahap pengembangan (*development*) ini akan dihasilkan asesmen autentik berupa instrumen penilaian yang dapat mengukur kompetensi sikap, pengetahuan, dan keterampilan. Kompetensi sikap diukur dengan teknik penilaian diri. Kompetensi pengetahuan diukur dengan teknik penilaian tes tertulis dengan bentuk uraian dan pilihan ganda. Kompetensi keterampilan diukur dengan teknik penilaian unjuk kerja dan produk. Indikator ketercapaian kompetensi disesuaikan dengan materi pembelajaran tentang sistem pernapasan. Kemudian asesmen autentik yang telah didesain divalidasi oleh ahli dan guru. Validasi dilakukan oleh 6 ahli dan 20 guru yang akan menilai sesuai atau tidak produk yang dihasilkan sebagai alat ukur pembelajaran. Selanjutnya akan diuji coba dan hasilnya menjadi bahan untuk merevisi asesmen autentik sebelum diimplementasikan dalam kegiatan pembelajaran.

Pada tahap implementasi (*implementation*), instrumen penilaian autentik yang sudah valid dan reliabel di gunakan untuk menilai pengetahuan, keterampilan dan sikap siswa. Instrumen ini akan diimplementasikan pada siswa kelas VIII SMPN 1 Warungkiara, dengan jumlah 39 siswa. Hasilnya kemudian diolah dan dianalisis untuk mengetahui tingkat pemahaman siswa.

Pada tahap evaluasi (*evaluation*), dilakukan evaluasi dengan menganalisis hasil uji coba lapangan. Lalu menyempurnakan produk berdasarkan hasil penelitian. Pemberian angket kepada siswa untuk mengetahui tanggapan siswa tentang asesmen autentik yang dikembangkan dan hasilnya dijadikan bahan perbaikan.

## **HASIL DAN PEMBAHASAN**

### **Hasil Validasi Ahli dan Guru IPA**

Validitas Isi terhadap instrumen penilaian kognitif, sikap, kinerja, dan produk meliputi validitas isi, konstruk, kebahasaan dan praktikalitas, dilakukan untuk mengetahui kelayakan instrumen tersebut. Validator dalam penelitian ini adalah enam orang ahli (dosen) dan 20 orang guru IPA. Data hasil validasi dianalisis menggunakan rumus CVR (*Content Validity Ratio*) dari Lawshe (1975) dan dilengkapi dengan tabel harga CVR kritis Lawshe (CVR kritis) menurut Wilson. Hasil yang diperoleh pada instrumen penilaian kognitif sebesar 0,99 dan 0,99. Sedangkan pada instrumen penilaian sikap hasil validitas isi oleh ahli sebesar 0.91 dan 0.97 hasil validitas isi oleh guru. Pada instrumen penilaian kinerja hasil validitas isi oleh ahli 0,99 dengan validitas isi oleh guru sebesar 0.94. Dan untuk instrumen penilaian produk hasil validitas isi oleh ahli sebesar 0,99 dengan validitas isi oleh guru 0.96.

### **Hasil Analisis Penilaian Kognitif**

Uji soal skala terbatas dilaksanakan pada satu sekolah, untuk menganalisis validitas soal, reliabilitas soal, daya pembeda soal, tingkat kesukaran soal pada bentuk soal pilihan ganda dan uraian. Validitas soal pilihan ganda pada skala terbatas sebanyak 30 soal dihitung dengan menggunakan rumus koefisien korelasi *produk moment* ( $r$ ). Jumlah seluruh siswa 37 orang, sehingga  $N=37$  dengan derajat kebebasan  $df = 35$ , nilai  $r$  tabel menunjukkan angka 0,325 dengan taraf signifikan 5%. Berdasarkan hasil analisis ke-30 butir soal uji soal skala terbatas diketahui bahwa soal yang masuk kategori valid berjumlah 21 butir soal (70%). Sedangkan soal yang masuk kategori tidak valid berjumlah 9 butir soal (30%). Soal-soal yang dinyatakan tidak valid karena memiliki nilai  $r$  hitung lebih kecil dari  $r$  tabel. Reliabilitas soal pilihan ganda dengan rumus Kuder Richardson (KR-20) sebesar 0,87 berarti tes yang digunakan memiliki reliabilitas yang sangat tinggi. Hasil analisis daya pembeda diketahui bahwa sebanyak 13 butir soal termasuk kategori sangat baik, 8 butir soal termasuk kategori baik, 1 butir soal

termasuk kategori cukup, dan 8 butir soal termasuk kategori jelek. Hasil analisis tingkat kesukaran diketahui bahwa sebanyak 8 butir soal termasuk kategori mudah, 19 butir soal termasuk kategori sedang, dan 2 butir soal termasuk kategori sukar.

Validitas soal uraian pada skala terbatas sebanyak 20 soal. Jumlah seluruh siswa 31 orang, sehingga  $N=31$  dengan derajat kebebasan  $df = 29$  nilai  $r$  tabel menunjukkan angka 0,355 dengan tarap signifikan 5%. Berdasarkan hasil analisis ke-20 butir soal, diketahui bahwa soal yang masuk kategori valid berjumlah 11 butir soal (55%). Sedangkan soal yang masuk kategori tidak valid berjumlah 9 butir soal (45%). Sembilan soal yang dinyatakan tidak valid tersebut karena memiliki nilai  $r$  hitung yang lebih kecil dari  $r$  tabel. Reliabilitas soal digunakan rumus Alpha Cronbach, sebesar 0,68 berarti tes yang digunakan memiliki reliabilitas tinggi. Daya pembeda soal diketahui sebanyak 3 soal kategori sangat baik, 8 soal kategori baik, 8 butir soal termasuk kategori cukup, dan 1 butir soal termasuk kategori jelek. Untuk analisis tingkat kesukaran diketahui sebanyak 6 butir soal termasuk kategori mudah, 11 butir soal termasuk kategori sedang, dan 3 butir soal termasuk kategori sukar.

Selanjutnya soal-soal yang valid pada skala terbatas diujicobakan pada skala luas. Validitas soal pilihan ganda skala luas sebanyak 21 soal, dengan jumlah siswa 89 orang, sehingga  $N=89$  dengan derajat kebebasan  $df = 87$ , nilai  $r$  tabel menunjukkan angka 0,208. Berdasarkan hasil analisis ke-21 butir soal, diketahui bahwa soal yang masuk kategori valid berjumlah 15 butir soal (71,4%). Sedangkan soal yang masuk kategori tidak valid berjumlah 6 butir soal (28,6%). Soal-soal yang dinyatakan tidak valid tersebut memiliki nilai  $r$  hitung yang lebih kecil dari  $r$  tabel. Reliabilitas soal sebesar 0,70 berarti tes yang digunakan memiliki reliabilitas tinggi. Analisis daya pembeda soal diketahui sebanyak 6 butir soal termasuk kategori sangat baik, 8 butir soal termasuk kategori baik, 2 butir soal termasuk kategori cukup, dan 5 butir soal termasuk kategori jelek. Analisis tingkat kesukaran diketahui sebanyak 2 butir soal termasuk kategori mudah, 16 butir soal termasuk kategori sedang, dan 3 butir soal termasuk kategori sukar.

Pada validitas soal uraian uji skala luas sebanyak 11 soal, diikuti 84 siswa, sehingga  $N=84$  dengan derajat kebebasan  $df = 82$ , nilai  $r$  tabel menunjukkan angka 0,214 dengan tarap signifikan 5%. Berdasarkan hasil analisis ke-11 butir soal uji diketahui bahwa soal yang masuk kategori valid berjumlah 8 butir soal (72,73%). Sedangkan soal yang masuk kategori tidak valid berjumlah 3 butir soal (27,27%). Lima soal yang dinyatakan tidak valid karena memiliki nilai  $r$  hitung yang lebih kecil dari  $r$  tabel. reliabilitasnya 0,66 berarti tes yang digunakan memiliki reliabilitas yang tinggi. Sedangkan analisis daya pembeda soal  $r$  soal uraian tersebut, diketahui bahwa sebanyak 7 butir soal termasuk kategori baik, 0

butir soal termasuk kategori cukup, dan 4 butir soal termasuk kategori jelek. Sedangkan analisis tingkat kesukaran soal diketahui sebanyak 2 butir soal termasuk kategori mudah, 9 butir soal termasuk kategori sedang, dan 0 butir soal termasuk kategori sukar.

#### **Hasil Analisis Penilaian Afektif**

Validitas pada penilaian skala sikap dihitung dengan menggunakan rumus koefisien korelasi produk moment. Jumlah seluruh siswa 116 orang, sehingga  $N=116$  dengan derajat kebebasan  $df = 114$ , nilai  $r$  tabel menunjukkan angka 0,18 dengan tarap signifikan 5% . Berdasarkan hasil analisis ke-30 butir pernyataan diketahui bahwa yang masuk kategori valid berjumlah 28 pernyataan (93,33%). Sedangkan yang masuk kategori tidak valid berjumlah 2 pernyataan (6,67%). Pernyataan yang dinyatakan tidak valid karena nilai  $r$  hitung lebih kecil dari  $r$  tabel. Hasil analisis reliabilitas pada instrument penilaian sikap dengan menggunakan rumus Alpha Cronbach, memiliki reliabilitas tinggi dengan nilai 0.76.

#### **Hasil Analisis Penilaian Psikomotor**

Validitas pada penilaian kinerja dihitung dengan menggunakan rumus koefisien korelasi produk moment. Jumlah seluruh siswa 40 orang, sehingga  $N=40$  dengan derajat kebebasan  $df = 38$ , nilai  $r$  tabel menunjukkan angka 0,31 dengan tarap signifikan 5% . Berdasarkan hasil analisis ke-tujuh aspek penilaian diketahui bahwa yang masuk kategori valid berjumlah 7 aspek penilaian (100%). Hal ini menunjukkan bahwa tujuh aspek penilaian tersebut memiliki nilai  $r$  hitung lebih besar dari  $r$  tabel. Hasil analisis reliabilitas pada instrument penilaian kinerja dengan menggunakan rumus Alpha Cronbach sebesar 0,53 hal ini nilai menunjukkan nilai reliabilitas yang cukup tinggi.

Validitas pada instrumen penilaian produk dihitung dengan menggunakan rumus koefisien korelasi produk moment. Jumlah seluruh siswa 40 orang, sehingga  $N=40$  dengan derajat kebebasan  $df = 38$ , nilai  $r$  tabel menunjukkan angka 0,31 dengan tarap signifikan 5% . Berdasarkan hasil analisis ke-lima aspek penilaian diketahui bahwa yang masuk kategori valid berjumlah 5 aspek penilaian (100%). Hal ini menunjukkan bahwa lima aspek penilaian tersebut memiliki nilai  $r$  hitung lebih besar dari  $r$  tabel. Reliabilitas pada instrumen penilaian produk dengan menggunakan rumus Alpha Cronbach sebesar 0,31 hal ini menunjukkan nilai reliabilitas yang rendah.

#### **Hasil Implementasi Asesmen Autentik**

Instrumen penilaian kognitif, afektif dan psikomotor yang telah memenuhi validitas, reliabilitas, daya pembeda dan tingkat kesukaran sesuai kemudian diterapkan pada satu kelas dengan jumlah siswa 39 orang,

Hasil penerapan instrumen asesmen autentik tersebut menunjukkan ketuntasan belajar pada aspek kognitif 82 %, aspek afektif 100 % dan aspek psikomotor rerata 87 %.

#### **Hasil Tanggapan Angket Siswa**

Hasil angket tanggapan siswa terhadap penilaian autentik yang terdiri dari 10 pertanyaan, direspon oleh 39 siswa, mendapatkan tanggapan sangat baik yaitu 91.02%. Sebanyak 9 pertanyaan pada skor diantara 76% sampai 100% dalam kriteria sangat baik dan satu pertanyaan pada kriteria baik diantara skor 51% sampai 75%.

### **PEMBAHASAN**

#### **Validasi Ahli dan Guru IPA**

Validator ahli yang memvalidasi instrumen penilaian kognitif, penilaian afektif, penilaian kinerja dan juga penilaian produk yang dikembangkan berjumlah enam orang dosen. Nilai CVR kritis untuk enam validator pada tingkat signifikansi 0,05 berdasarkan tabel schipper adalah 0,672 (Wilson, 2012). Sedangkan validasi yang dilakukan oleh 20 guru IPA pada instrumen penilaian kognitif, penilaian sikap, penilaian kinerja dan penilaian produk, nilai CVR kritisnya 0,368 . Instrumen penilaian dikatakan valid apabila nilai CVR hitung yang diperoleh lebih tinggi dari pada nilai CVR kritis. Apabila dibandingkan dengan nilai kritisnya maka nilai CVR lebih besar daripada nilai kritisnya. Dengan demikian Instrumen Asesmen Autentik yang dikembangkan valid atau layak dipergunakan. Hal ini sesuai dengan hasil penelitian Basit dan Wardana (2016) bahwa hasil uji validasi para ahli menyatakan sangat valid.

#### **Validitas Butir Soal Instrumen Asesmen Autentik**

Hasil penelitian instrumen asesmen autentik pada materi sistem pernapasan secara umum telah memiliki validitas yang bagus. Sebagian besar menunjukkan butir soal atau aspek penilaian yang masuk dalam kriteria valid dan sebagian kecil masuk kriteria tidak valid. Sejalan dengan hasil penelitian Kustitik dan Hadi (2016) bahwa perangkat penilaian autentik yang dikembangkan dinyatakan memenuhi kriteria valid, reliabel, dan praktis. Validitas item merupakan validitas yang dimiliki oleh masing-masing butir item yang membangun tes tersebut. Validitas item dapat diketahui dengan melihat besar kecilnya dukungan yang diberikan oleh masing-masing butir item yang bersangkutan terhadap tes sebagai keseluruhan. item soal yang tidak valid berarti mengindikasikan bahwa item soal tersebut tidak mewakili tes secara keseluruhan atau item tes mengindikasikan tidak mengukur materi atau level materi yang sama dibanding item soal lainnya. Kemungkinan tidak validnya suatu item soal karena petunjuk tes yang tidak jelas, kata-kata yang digunakan kurang dipahami, waktu pengerjaan yang kurang tepat,



jawaban gampang diprediksi siswa, adanya kecurangan dalam pelaksanaan tes. Butir soal yang tidak valid tersebut sebaiknya direvisi dan butir soal yang valid dapat digunakan kembali untuk tes. Hasil penelitian ini sesuai dengan teori validitas menurut Sudijono, A. (2006) bahwa butir soal yang memiliki validitas tinggi mencerminkan soal tersebut telah memiliki kehandalan dan tidak perlu diragukan lagi ketepatannya dalam mengukur kemampuan peserta didik. Alat ukur yang baik harus memenuhi syarat validitas. Suatu tes dikatakan valid jika dapat mengukur apa yang seharusnya diukur. Untuk butir soal yang memiliki validitas rendah mencerminkan soal tersebut tidak valid sehingga perlu dilakukan tindakan terhadap soal tersebut.

Berdasarkan uraian di atas dapat disimpulkan bahwa sebagian besar soal termasuk soal yang cukup baik berdasarkan validitasnya. Untuk butir soal yang tidak valid sebaiknya diadakan revisi dengan cara meningkatkan penguasaan teknis tentang cara-cara pembuatan soal tes. Soal dapat menjadi valid karena konstruksinya baik dan mencakup materi yang benar-benar mencakup keseluruhan yang akan diukur.

### **Reliabilitas**

Reliabilitas merupakan konsistensi suatu alat ukur dalam pengukuran. Selain syarat validitas, alat ukur yang baik juga harus memenuhi syarat reliabilitas. Reliabilitas menunjukkan suatu pengukuran yang dapat diandalkan atau dapat dipercaya, artinya hasil yang dicapai oleh tes itu konsisten atau ajeg, tidak menunjukkan perubahan-perubahan yang berarti. Menurut Anas Sudijono (2006) reliabilitas menunjukkan sejauhmana hasil pengukuran dengan alat tersebut dapat dipercaya, ajeg dan stabil. Reliabilitas juga diartikan sebagai tingkat kepercayaan hasil suatu pengukuran, (Azwar, 2016).

Hasil analisis reliabilitas pada instrumen penilaian kognitif, afektif dan psikomotor menunjukkan kategori sangat tinggi, sampai rendah. Tinggi rendahnya nilai reliabilitas dapat dipengaruhi oleh sebaran skor dalam kelompok siswa yang diukur, tingkat kesulitan tes, dan jumlah butir tes. Semakin besar penyebaran skor semakin tinggi koefisien korelasinya, soal-soal yang tingkat kesulitannya sedang, reliabilitasnya tinggi dan semakin banyak jumlah butir soal semakin tinggi reliabilitasnya.

### **Daya Pembeda**

Daya pembeda adalah ukuran kekuatan pembeda suatu butir soal. Daya pembeda berkaitan dengan kemampuan butir soal untuk membedakan antara siswa yang menguasai materi dengan peserta didik yang kurang menguasai materi. Nilai daya beda yang positif berarti memiliki daya beda tinggi sedangkan nilai daya beda yang negatif berarti memiliki daya beda yang rendah. Pada

analisis yang dilakukan secara keseluruhan daya beda soal diketahui ada beberapa soal yang memiliki daya beda jelek atau yang bernilai antara 0,0 - 0,20.

Soal pada kriteria diterima atau baik berarti soal tersebut mempunyai kemampuan untuk membedakan antara siswa yang sudah memahami materi dan siswa yang belum memahami materi. Soal pada kriteria tidak diterima/tidak dapat dipakai maka harus dibuang karena tidak dapat membedakan antara siswa yang sudah memahami materi dan siswa yang belum memahami materi. Suatu soal tidak dapat membedakan antara siswa yang memahami materi dan siswa yang belum memahami materi dimungkinkan karena kunci jawaban soal tidak tepat, butir soal mempunyai dua atau lebih kunci jawaban, kompetensi yang diukur tidak jelas, pengecoh tidak berfungsi, dan materi yang ditanyakan terlalu sulit.

Butir soal yang memiliki daya pembeda yang baik sebaiknya dimasukkan ke dalam bank soal dan dapat digunakan karena kualitasnya memadai. Butir soal yang memiliki daya pembeda masih rendah, sebaiknya direvisi sehingga dapat digunakan lagi pada tes berikutnya atau dibuang dan tidak digunakan lagi untuk mengevaluasi hasil belajar siswa. Butir soal yang memiliki daya pembeda negatif sebaiknya pada tes selanjutnya tidak digunakan lagi/dibuang karena butir soal yang demikian kualitasnya sangat jelek sehingga lebih banyak siswa pandai yang menjawab salah daripada siswa yang kurang pandai.

### **Tingkat Kesukaran**

Tingkat kesukaran ukuran tingkat kesulitan atau kemudahan suatu butir soal. Butir soal yang baik apabila tingkat kesukarannya tidak terlalu sulit dan juga tidak terlalu mudah, yaitu kategori sedang dengan indeks tingkat kesukaran 0,31-0,70. Butir soal yang terlalu sulit akan menyebabkan siswa putus asa dikarenakan siswa merasa tidak mampu untuk mengerjakan soal tersebut. Butir soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usahanya untuk memecahkan masalah. Butir soal yang terlalu mudah juga tidak baik karena kurang mampu membedakan siswa yang pandai dan siswa yang kurang pandai. Hasil analisis untuk butir soal yang termasuk kategori baik, artinya tingkat kesukarannya sedang, sebaiknya butir soal tersebut dimasukkan ke dalam bank soal, sehingga dapat digunakan lagi pada tes yang akan datang. Untuk butir soal yang termasuk kategori sukar, butir soal tersebut dibuang dan tidak dikeluarkan lagi pada tes yang akan datang atau diteliti ulang faktor-faktor yang menyebabkan butir soal tersebut sulit dijawab benar oleh siswa dari segi kejelasan kalimat, kejelasan petunjuk pengerjaan, kemungkinan adanya istilah yang sulit dipahami, ataupun kesalahan pemahaman materi ajar. Butir soal yang termasuk kategori sukar bukan berarti tidak diperlukan dalam evaluasi hasil belajar, namun butir soal yang termasuk kategori sukar harus disesuaikan dengan proporsi yang ideal di dalam sebuah tes yaitu sekitar 20% dari jumlah keseluruhan soal yang

diujikan. Untuk butir soal yang termasuk kategori mudah, tidak digunakan lagi pada tes yang akan datang atau diteliti ulang untuk mengetahui faktor-faktor yang menyebabkan butir soal begitu mudah dijawab dengan benar oleh siswa. rata-rata soal dengan kategori mudah, pengecohnya tidak berfungsi dengan baik sehingga siswa mudah mengetahui mana opsi yang merupakan kunci dan mana yang merupakan pengecoh. Oleh sebab itu dapat dilakukan revisi dengan mengganti pilihan jawaban. Setelah dilakukan perbaikan, butir soal yang bersangkutan masih harus diujikan lagi untuk mengetahui apakah derajat kesukaran item menjadi lebih baik daripada sebelumnya. Seperti soal pada kategori sukar, soal dengan kategori mudah juga diperlukan dalam evaluasi hasil belajar siswa namun dengan proporsi yang ideal yaitu sekitar 30% dari keseluruhan soal.

### **Implementasi Asesmen Autentik**

Keberhasilan siswa dalam belajar dapat dilihat dari pencapaian hasil belajar yang diperoleh, baik aspek kognitif, afektif dan psikomotor. Kriteria Ketuntasan Minimal adalah kriteria ketuntasan belajar yang ditentukan oleh satuan pendidikan melalui prosedur tertentu, jika hasil belajar yang diperoleh siswa sama dengan atau melampaui KKM berarti siswa tersebut telah tuntas dalam kompetensi yang telah ditentukan. Dari hasil implementasi instrumen penilaian kognitif, afektif dan psikomotor dapat disimpulkan bahwa ketuntasan belajar yang dicapai siswa lebih dari 75 %. Seperti yang dikemukakan Kunandar (2014), tujuan analisis tingkat ketuntasan pencapaian kompetensi siswa adalah untuk memetakan banyaknya siswa yang sudah dan belum menguasai kompetensi yang ditentukan.

Penilaian hasil belajar yang fungsional harus memenuhi syarat-syarat tertentu, antara lain instrumen atau alat ukur yang digunakan harus valid dan reliabel, artinya dari segi penyusunan telah memenuhi kaidah-kaidah penulisan soal, baik dari aspek konstruksi maupun materi. Dengan demikian penilaian yang valid dan reliabel akan menghasilkan informasi tingkat penguasaan kompetensi siswa yang akurat dan terpercaya. Menurut Kunandar (2014), penyusunan instrumen penilaian harus sesuai dengan karakteristik dan kompleksitas materi yang terdapat dalam Standar Isi dan Standar Kompetensi Lulusan. Dengan demikian instrumen tersebut akan memberikan informasi yang akurat tentang tingkat pencapaian kompetensi peserta didik. Hal ini sejalan dengan hasil penelitian yang dikemukakan Ermawati dan Hidayat (2017) bahwa metode penilaian autentik diyakini sangat tepat untuk diimplementasikan dalam proses pembelajaran, karena penilaian ini memiliki banyak dampak positif, terutama bagi siswa. Meskipun teknik penilaian autentik ini lebih fokus pada peningkatan kualitas hasil belajar siswa daripada prestasi belajar siswa.

### **Angket Tanggapan Siswa terhadap Asesmen Autentik**

Pertanyaan pertama 100% siswa menyetujui cara pengambilan nilai yang dilakukan, pertanyaan ke dua 69.2% siswa menyatakan soal-soal yang diberikan dapat dikerjakan dengan mudah, pertanyaan ke tiga 82.1% siswa dapat memahami soal dari segi bahasa dengan mudah, pertanyaan ke empat 79.5% siswa tidak merasa kesulitan mengerjakan soal yang diberikan, pertanyaan ke lima 84.6% siswa menyatakan cukup waktu ketika mengerjakan soal yang diberikan, pertanyaan ke enam 97.4% siswa menyatakan bahwa soal yang diberikan dapat mengungkap kemampuan mereka dalam belajar, pertanyaan ke tujuh 97.4% siswa menyetujui penilaian sikap yang dilakukan, pertanyaan ke delapan 100% siswa menyetujui penilaian kinerja yang dilakukan, pertanyaan ke sembilan 100% siswa menyetujui penilaian produk berupa poster yang dilakukan, dan pertanyaan terakhir 100% siswa menyatakan bahwa asesmen autentik bermamfaat/berguna untuk memotivasi mereka, dalam meningkatkan hasil belajar. Sehingga secara keseluruhan tanggapan siswa terhadap asesmen autentik sangat baik, dengan kata lain siswa menyetujui adanya asesmen autentik yang dilaksanakan dalam pembelajaran. Seperti yang dikemukakan Astuti, dkk (2012) dari hasil penelitiannya bahwa kepraktisan instrumen asesmen ditunjukkan oleh adanya respon positif siswa.

### **PENUTUP**

Penelitian pengembangan asesmen autentik yang dilakukan menghasilkan produk yang meliputi instrument penilaian kognitif berupa soal uraian dan pilihan ganda pada materi sistem pernapasan pada manusia kelas 8 Sekolah Menengah Pertama, instrumen penilaian sikap berupa penilaian diri, instrumen penilaian keterampilan berupa penilaian kinerja dan produk. Adapun kesimpulan hasil pengembangan instrumen asesmen autentik ini yaitu instrumen penilaian kognitif yang telah dikembangkan untuk materi sistem pernapasan pada manusia kelas VIII, Sekolah Menengah Pertama. Dengan ranah kognitif mulai dari memahami (C2), mengaplikasikan (C3), menganalisis (C4), mengevaluasi (C5), mencipta (C6). Untuk instrumen penilaian sikap dan keterampilan memiliki profil 1) konten yang sesuai dengan indikator penilaian. 2) praktis karena memudahkan dalam pengukuran dan penilaian. 3) berguna untuk memotivasi siswa dalam meningkatkan hasil belajar.

Keseluruhan produk hasil pengembangan dinyatakan layak untuk digunakan berdasarkan hasil validasi ahli dan guru IPA, karena nilai CVR hitung ahli dan guru yang dicapai 0,91 sampai 0,99 sedangkan nilai CVR kritisnya 0,672 dan 0,368. Dengan demikian CVR hitung yang diperoleh lebih tinggi dari pada nilai CVR kritis. Hasil analisis validitas item soal produk yang dikembangkan 55 % sampai 100 % valid dengan tingkat reliabilitas sangat tinggi sampai rendah.

Dengan demikian berdasarkan validitas, reliabilitas, daya pembeda dan tingkat kesukaran pada instrumen penilaian kognitif terdapat 15 soal pilihan ganda dan 8 soal uraian valid dengan reliabilitas tinggi sehingga layak digunakan. Untuk instrumen penilaian sikap terdapat 28 pernyataan valid dengan reliabilitas tinggi yang dapat dipergunakan. Sedangkan pada instrumen penilaian kinerja 7 aspek penilaian yang valid dengan reliabilitas cukup tinggi dapat digunakan dan pada instrumen penilaian produk 5 aspek penilaian valid dengan tingkat reliabilitas rendah, dapat digunakan serta penerapan instrumen penilaian kognitif, afektif dan psikomotor yang valid dan reliabel dapat mengukur tingkat pencapaian ketuntasan belajar.

Berdasarkan penelitian tentang pengembangan instrumen asesmen autentik yang telah dilakukan, beberapa saran yang dapat disampaikan yaitu produk hasil penelitiang pengembangan asesmen autentik ini dapat digunakan oleh guru-guru dalam pelaksanaan penilaian, agar hasil penilaian dapat mencapai kompetensi ranah kognitif, afektif dan psikomotor secara nyata (autentik). Peneliti berikutnya dapat mengembangkan produk pada Kompetensi Dasar lainnya. Implementasi produk pengembangan asesmen autentik ini perlu dilakukan pada skala yang luas.

## **REFERENSI**

- Astuti, W. Puji, Prasetyo A. P. B, & Rahayu E. S. (2012). Pengembangan Instrumen Asesmen Autentik Berbasis Literasi. Sains Pada Materi Sistem Ekskresi. *Journal Lembaran Ilmu Kependidikan*, 41 (1), 39-43.
- Atac, B. A. (2012). Foreign Language Teachers' Attitude toward Authentic Assessment in Language. *The Journal of Language and Linguistic Studies*, 8 (2), 7-17.
- Azwar, Saifuddin. (2016). *Sikap Manusia: Teori dan Pengukurannya (edisi 2)*. Yogyakarta: Pustaka Pelajar.
- Basit, A & Wardana L. A. (2016). Pengembangan Perangkat Asesmen Autentik dalam Pembelajaran Pendidikan Kewarganegaraan. *Jurnal Ilmiah Pendidikan Pancasila dan Kewarganegaraan*, 1 (2), 73-78.
- Boadu K., Anthony B., Isaac E, Thoephilus K. B., Fredrick Y. K. (2014). Teachers' Perception Of Authentic Assessment Techniques Practice In Social Studies Lessons In Senior High Schools In Ghana. *Journal of Penelitian Pendidikan dan Ilmu Informasi*, 1 (4), 62-68.
- Brown, H. D. (2004). *Language Asesmen : Principles and Classroom Practice. White Planis*. New York: Pearson Education.
- Ermawati, S & Hidayat, T. (2017). Penilaian Autentik dan Relevansinya dengan Kualitas Hasil Pembelajaran (Persepsi Dosen dan Mahasiswa IKIP PGRI Bojonegoro). *Jurnal Pendidikan Ilmu Sosial*, 27 (1), 92-103.
- Gulikers, Judith T. M., Theo J. Bastiaens & Paul Kirschner. (2004). A five Dimensional Framework for Authentic Assessment. *ETR&D*, 52 (3), 67-86.

- Johnson. (2002). *Cooperative Learning Methode: A. Meta-Analysis*.
- Kunandar. (2014). *Penilaian Autentik (Penilaian Hasil Belajar Peserta Didik berdasarkan Kurikulum 2013) Suatu Pendekatan Praktis Edisi Revisi*. Jakarta: PT. Raja Grafindo Persada
- Kustitik & Hadi S. (2016). Pengembangan Perangkat Penilaian Autentik Mata Pelajaran Prakarya Dan Kewirausahaan Di SMK. *Jurnal Pendidikan Vokasi*, 6 (2), 184-197.
- Lawshe, C. H. (1975). A Quantitative Approach To Content Validity. *Personnel. Psychology*, 28, 563-575.
- Liandari, E. Parsaoran S, Ida K. & Isnaini. (2017). Upaya Meningkatkan Kemampuan Merumuskan dan Menguji Hipotesis melalui Pendekatan Keterampilan Proses Sains dengan Metode Praktikum. *Jurnal Wahana Pendidikan Fisika*, 2 (1), 50-55
- Nurgiyantoro, B. (2008). Asesmen Autentik. *Cakrawala Pendidikan*, 27 (3), 250-261.
- Permendikbud. (2014). *Penilaian Hasil Belajar oleh Pendidik pada Pendidikan dasar dan pendidikan Menengah*. Jakarta : Kementerian Pendidikan dan Kebudayaan.
- Rizal, Muhammad. (2014). Pengaruh Pembelajaran Inkuiri Terbimbing dengan Multi Representasi terhadap Keterampilan Proses Sains dan Penguasaan Konsep IPA Siswa SMP. *Jurnal Pendidikan Sains*, 2 (3), 159-165.
- Rukmana, T dan Mundilarto. (2016). Keterlaksanaan Penilaian Autentik Mata Pelajaran Fisika SMA Negeri. *Jurnal Inovasi Pendidikan IPA*, 2 (1), 111-121.
- Sudijono, A. (2006). *Pengantar Evaluasi Pendidikan*. Jakarta: RajaGrafindo Persada.
- Wilson, F. Robert, Pan,Wei and Schumsky, Donald A.. (2012). Recalculation of the Critical Values for Lawshe's Content Validity Ratio. *Measurement and Evaluation in Counseling and Development*, 45, 197-210 [Originally Published Online] DOI: 10.1177/0748175612440286