

Pengembangan Instrumen Tes *Two-Tier Higher Order Thinking Skills (HOTS)* Berbasis IBT pada Materi Titrasi Asam Basa Untuk Siswa SMA/MA

Development of IBT Based Two-Tier Higher Order Thinking Skills (HOTS) Test Instruments on Acid-Base Titration Materials for SMA/MA Students

Shiva Atika Suri¹, and Andromeda^{1*}

¹Program Studi Pendidikan Kimia, Fakultas Matematika dan IPA, Universitas Negeri Padang, Sumatera Barat, Indonesia.

andromeda@fmipa.unp.ac.id

ABSTRACT

This research aims to produce a valid and reliable IBT based two-tier Higher Order Thinking Skills (HOTS) test instrument on acid-base titration material. This is a research and development (R&D) which conducted by the Treagust model which has 3 stages, namely determining content, obtaining information about students regarding the instrument developed, and developing a diagnostic test. The validity was processed by Content validity analysis using the CVR method and construct validity using Aiken V. The CVR value of content validity consisting of questions' stimulus, questions, key answer and reasons is in a row of 0,97, 1, 1, 1. The results of content validity of propositional knowledge statement and concept maps are 1, and the results of validity construct consisting of material aspects, presentation, language, and other rules in a row of 0.95, 0.95, 0.97, 0.98. The reliability of instrument test for first tier and second tier respectively namely 0,89 and 0,90. The results show that the developed test instrument is valid and has good item quality.

Keywords: Instrument Tes, Two-Tier, Higher Order Thinking Skills, Acid-Base Titration

ABSTRAK

Penelitian ini merupakan penelitian pengembangan yang bertujuan untuk menghasilkan instrument tes *two-tier Higher Order Thinking Skills (HOTS)* berbasis IBT pada materi titrasi asam basa yang valid dan reliabel. Penelitian ini merupakan penelitian *research and development (R&D)* dengan menggunakan model Treagust yang mempunyai 3 tahapan yaitu penentuan isi, mendapatkan informasi mengenai peserta didik terkait instrumen yang dikembangkan, dan mengembangkan tes diagnostik. Teknik analisis validitas isi menggunakan metode CVR sedangkan validitas konstruk menggunakan Aiken V. Berdasarkan hasil penelitian, nilai CVR validitas isi yang terdiri dari stimulus soal, pertanyaan soal, jawaban soal, serta alasan soal berturut-turut sebesar 0,97, 1, 1, 1, dan validitas isi berupa pernyataan pengetahuan preposisi dan peta konsep yaitu 1. Nilai validitas konstruk yang terdiri atas aspek materi, penyajian, Bahasa, dan aturan tambahan berturut-turut sebesar 0.95, 0.95, 0.97, 0.98. Instrumen tes memiliki tingkat reliabilitas *first tier* 0,89 dan *second tier* 0,90. Hasil menunjukkan bahwa instrumen tes yang dikembangkan sudah valid dan memiliki kualitas instrumen tes yang baik.

Kata Kunci: Instrumen Tes, Two-Tier, Higher Order Thinking Skills, Titrasi Asam Basa

PENDAHULUAN

Pada abad 21, perkembangan teknologi dan informasi berjalan sangat cepat. Perkembangan pendidikan dan budaya ditingkat internasional menjadikan hal tersebut sebagai tantangan tersendiri yang harus dihadapi. Para peserta didik harus melek terhadap informasi, media dan teknologi. Kata “melek” bukan hanya sekedar tahu dan bisa menggunakan, akan tetapi dengan kemampuan berpikir tingkat tinggi, peserta didik dapat mengeksplor ide dan pemikirannya sehingga informasi, media, dan teknologi yang sudah ada dapat diaplikasikan dengan baik.

Kurikulum 2013 menerapkan penilaian autentik untuk menilai kemajuan belajar peserta didik yang meliputi sikap, pengetahuan, dan keterampilan (Kemendikbud, 2017). Penilaian merupakan rangkaian kegiatan agar memperoleh, menganalisis, serta menafsirkan data tentang proses dan hasil belajar peserta didik yang dilakukan secara sistematis dan berkesinambungan sehingga menjadi informasi yang bermakna dalam pengambilan keputusan (Purwanto, 2011).

Penilaian hasil belajar diharapkan dapat membantu peserta didik untuk meningkatkan keterampilan berpikir tingkat tinggi *Higher Order Thinking Skills* (HOTS), karena keterampilan berpikir tingkat tinggi dapat mendorong peserta didik untuk mempunyai bekal dalam sejumlah kompetensi yang di perlukan untuk menghadapi abad ke-21. Beberapa kompetensi tersebut yaitu 4C terdiri atas (1) kemampuan berpikir kritis; (2) kreativitas; (3) kerja sama; (4) kemampuan berkomunikasi (Partinem, 2019).

Menurut Thomas & Thorne (2009) HOTS merupakan cara berpikir yang lebih

tinggi dari pada menghafalkan fakta, mengemukakan fakta, atau menerapkan peraturan, rumus dan prosedur (Nugroho, 2021). Kemampuan berpikir tingkat tinggi merupakan salah satu kemampuan yang harus dimiliki peserta didik untuk menghadapi perkembangan zaman dan persaingan global. Oleh sebab itu, pembelajaran disekolah terutama SMA memerlukan proses yang melibatkan kemampuan berpikir tingkat tinggi untuk dinilai dan di evaluasi agar pencapaian peserta didik dapat diketahui. Penilaian yang dilakukan harus mengacu pada pencapaian Kompetensi Dasar (KD). Level kognitif dalam KD terbagi menjadi LOTS dan HOTS. LOTS terdiri dari mengingat (C1), memahami (C2), mengaplikasikan (C3), sedangkan HOTS terdiri dari menganalisis (C4), mengevaluasi (C5), dan mengkreasi (C6) (Anderson & Krothwhl, 2010).

Programme for International Students Assesment (PISA) merupakan penilaian internasional untuk menguji kemampuan akademis dari peserta didik di dunia yang diuji setiap 3 tahun sekali. Berdasarkan laporan hasil PISA, dua tahun terakhir yaitu tahun 2015 dan 2018, Indonesia menempati posisi yang tergolong rendah, dimana tahun 2015 Indonesia berada pada peringkat 64 dari 72 negara, tahun 2018 Indonesia berada pada peringkat 69 dari 76 negara. Hal ini menampilkan literasi sains peserta didik masih rendah. Keahlian berpikir masih hanya cenderung mengingat, melaporkan kembali, ataupun merujuk tanpa melaksanakan pengolahan (Nugroho, 2021).

Berdasarkan hasil angket yang diberikan melalui *google form* kepada 8 guru kimia di sekolah yang berbeda,

diperoleh 62,5% guru menggunakan instrument tes LOTS. Hal ini dikarenakan sumber instrumen tes berasal dari buku cetak dan LKS yang sebagian besar hanya memuat soal LOTS. Selain itu dengan perkembangan teknologi dan informasi, pada saat ini tes dapat dilakukan dengan menggunakan internet yang ada dengan berbagai *platform* atau media salah satunya dengan *google form*.

Menurut Maulidiansyah (2018) kelebihan dari *Google Form* yaitu pembuatannya mudah, gratis, jumlah survey yang dibuat tidak terbatas, jumlah responden yang tidak terbatas, jawaban survey dan data dikumpulkan otomatis di spreadsheet, dapat memasukkan video atau gambar, dan survey dapat disebarakan lewat media sosial. Penilaian dengan berbasis internet atau IBT di dukung dengan kondisi yang terjadi pada saat ini, yang dimana dunia sedang dilanda dengan musibah COVID-19. Akibat dari pandemic ini, proses belajar mengajar di lakukan secara daring. Dengan pembelajaran daring, penilaian pendidikan juga dilakukan secara daring. Oleh karena itu guru membutuhkan instrument tes *two-tier* HOTS berbasis internet salah satunya pada materi titrasi asam basa.

Two-tier multiple choice merupakan soal pilihan ganda bertingkat dua. Tingkat pertama berisi pertanyaan mengenai konsep yang disajikan selanjutnya tingkat kedua berisi tentang alasan peserta didik dalam memilih jawaban pada pertanyaan di tingkat pertama (Tuysuz, 2009). Berdasarkan uraian latar belakang, maka dibutuhkan instrumen tes *two-tier Higher Order Thinkig Skills* (HOTS) berbasis IBT pada materi titrasi asam basa yang valid dan reliabel.

METODE

Penelitian pengembangan atau *Research and Development* (R&D) merupakan metode yang digunakan dalam penelitian ini. Penelitian ini menghasilkan instrumen tes *two-tier Higher Order Thinking Skills* (HOTS) berbasis IBT pada materi titrasi asam basa. Model pengembangan yang dilakukan dalam penelitian ini adalah model pengembangan Treagust yang dikembangkan oleh David Treagust. Model ini memiliki 3 tahapan utama yaitu penentuan isi, mendapatkan informasi mengenai peserta didik terkait instrumen yang dikembangkan dan yang terakhir adalah mengembangkan tes diagnostic (Treagust, 1988).

Penelitian dilakukan di SMAN 12 Padang. Penelitian ini berlangsung pada semester ganjil tepatnya pada bulan juli sampai Desember pada tahun ajaran 2021/2022. Objek dari penelitian ini merupakan kumpulan instrumen tes *two-tier Higher Order Thinking Skills* (HOTS) berbasis IBT pada materi titrasi asam basa untuk siswa SMA/MA. Subyek dari peneltian ini yaitu Dosen jurusan kimia FMIPA UNP, guru kimia SMAN 12 Padang dan SMAN 1 Lubuk Alung.

Data dikumpulkan melalui tes dan angket validasi. Pada lembar angket validasi isi menggunakan formula CVR dengan jawaban berikut :

Tabel 1. Nilai Angket Validasi Isi

Nilai	Kategori
1	Ya
0	Tidak

Metode yang di pakai untuk menganalisis validitas isi adalah Lawshe's CVR (*Content Validity Ratio*). Formula yang diajukan oleh Lawshe adalah :

$$CVR = \frac{N - \frac{N}{2}}{\frac{N}{2}}$$

Keterangan :

CVR = nilai conten validity ratio

n = banyak ahli yang menjawab valid

N = jumlah total ahli

(Lawshe, 1985)

Item soal dapat diterima apabila nilai CVR pada item soal tersebut sama atau lebih besar dari nilai kritis, sedangkan item soal akan ditolak apabila nilai CVR item soal tersebut kurang dari nilai kritis dalam jumlah ahli. Jika terdapat 6 orang ahli maka nilai kritis CVR adalah 0,672 (Wilson dkk., 2012). Item soal memiliki kevalidan yang cukup apabila jika lebih dari setengah jumlah ahli memberikan jawaban valid.

Teknik analisis validitas konstruk diukur dengan menggunakan penyajian item soal yang terdiri dari aspek materi, penyajian, bahasa, dan aturan tambahan. Analisis dilakukan dengan memberikan angket yang berisikan pertanyaan kemudian validator memberikan kritik dan saran terhadap masing-masing pertanyaan tersebut. Penilaian validator terhadap masing-masing pertanyaan dianalisis menggunakan *Aiken's V*, dengan rumus sebagai berikut :

$$V = \frac{\sum s}{n(c-1)}$$

$$s = r - l_0$$

Keterangan :

v = indeks kesepakatan rater

l_0 = angka penilaian validitas yang terendah (dalam hal ini = 1)

c = angka penilaian validitas yang tertinggi (dalam hal ini = 5)

r = angka yang diberikan oleh seorang penilai

n = jumlah rater

Uji validitas konstruk menggunakan angket dengan mengaplikasikan indeks Aiken V yang terdiri atas 4 tingkatan diantaranya yaitu : Tidak Relevan (TR), Kurang Relevan (KR), Cukup Relevan (CR), dan Relevan (R) yang selanjutnya diolah dan dijelaskan sesuai dengan skala Aiken V. Indeks V yang diperoleh dari hasil perhitungan dapat diterima kevalidannya jika memenuhi nilai minimum kevalidan berdasarkan koefisien validitas pada tabel yang diberikan Aiken (Aiken, 1985).

Setelah dilakukan validasi, selanjutnya menguji kualitas dari instrument tes. Kualitas tes diperoleh dengan melakukan uji coba kepada 20 orang peserta didik SMAN 12 Padang. Tes dilakukan dengan menggunakan media *google form*. Peserta didik menjawab soal dengan memilih satu pertanyaan yang dianggap benar terhadap lima pilihan jawaban lainnya. Kualitas tes di analisis dengan menggunakan aplikasi ANATES versi 4.0.9. Aplikasi ini dapat menganalisis reliabilitas, indeks kesukaran soal, daya pembeda butir soal, pada instrumen tes.

HASIL DAN DISKUSI

Instrumen tes two-tier *Higher Order Thinking Skills* (HOTS) berbasis IBT pada materi titrasi asam basa dikembangkan dengan melalui tiga tahapan utama yaitu penentuan isi, mendapatkan informasi mengenai peserta didik terkait instrumen yang dikembangkan, dan yang terakhir adalah mengembangkan tes diagnostik.

Penentuan Isi

Tahap penentuan isi terdapat beberapa langkah kegiatan yang dilakukan oleh peneliti yang diantaranya yaitu mengidentifikasi pernyataan pengetahuan proposisi, mengembangkan peta konsep dan menghubungkan pengetahuan proposisi dengan peta konsep, melakukan validasi isi pernyataan pengetahuan preposisi dengan peta konsep. Setelah mengkaji IPK titrasi asam basa didapatkan sebanyak 11 pernyataan pengetahuan proposisi pada materi titrasi asam basa.

Pada analisis pernyataan pengetahuan proposisi materi titrasi asam basa, diperoleh 4 konsep prasyarat yang diperlukan sebelum mempelajari materi senyawa titrasi asam basa. Konsep prasyarat 4 terkait stoikiometri, konsep prasyarat 3 terkait reaksi kimia, selanjutnya konsep prasyarat 2 terkait asam basa, dan konsep prasyarat 1 terkait hidrolisis garam dan larutan penyangga.

Peta konsep yang peneliti kembangkan untuk penelitian ini didasarkan pada pernyataan pengetahuan proposisi yang telah dibuat. Peta konsep yang telah dibuat ini disesuaikan hirarkinya berdasarkan susunan konsep-konsep prasyarat hingga topik titrasi asam basa. Selanjutnya tahap menghubungkan pengetahuan proposisi dengan peta konsep. Tujuan dari tahap ini yaitu untuk melihat hubungan yang tepat antara peta konsep dan pengetahuan pernyataan proposisi yang telah dibuat sehingga dapat dilakukan validasi isi oleh para ahli. Didapatkan bahwa terdapat kecocokan antara pengetahuan proposisi dan peta konsep.

Validitas isi pernyataan pengetahuan proposisi dengan peta konsep sangat ditekankan guna memperoleh kualitas instrumen penelitian yang baik. Hasil data analisis isi dari para ahli yaitu dosen dan guru kimia menggunakan formula CVR dengan hasil 1 sehingga dapat disimpulkan

bahwa pernyataan pengetahuan preposisi dengan peta konsep sudah valid.

Mendapatkan Informasi Mengenai Berpikir Tingkat Tinggi Peserta Didik

Tahapan ini bertujuan untuk mengetahui tingkat berpikir peserta didik yang juga dipelajari oleh banyak peneliti lain terkait dengan identifikasi dalam penerapan instrument tes HOTS pada peserta didik. Tahapan ini dilakukan dengan cara pemeriksaan literatur yang relevan. Peneliti mendapatkan informasi mengenai berpikir tingkat tinggi pada peserta didik dengan sumber pustaka berjumlah 7 jurnal.

Berdasarkan studi literatur yang telah dilakukan, di dapatkan informasi bahwa salah satu cara untuk mengukur berpikir tingkat tinggi peserta didik dengan mengembangkan instrumen tes *two-tier multiple choice* berbasis *Higher Order Thinking Skills* (HOTS).

Mengembangkan Tes Diagnostik

Adapun Langkah yang dilakukan untuk mengembangkan instrumen tes diantaranya yaitu: merancang kisi-kisi, mengembangkan tes diagnostik, melakukan perbaikan.

Merancang kisi-kisi

Kisi-kisi dirancang untuk memudahkan dalam mengembangkan instrumen tes. Kisi-kisi disesuaikan dengan kompetensi dasar titrasi asam basa. Kisi-kisi ini memuat IPK, indikator soal, materi, level kognitif, bentuk soal yang telah sesuai dengan pernyataan preposisi dan peta konsep.

Mengembangkan tes diagnostic

Pengembangan instrumen tes ini peneliti merumuskan sebanyak 25 soal dengan jenis soal *two-tier multiple choice* dilengkapi lima pilihan jawaban. Instrumen ditulis berdasarkan kisi-kisi soal yang telah dirumuskan. Butir soal mengandung

stimulus soal yang merupakan ciri khas dari instrumen tes HOTS.

Butir Soal :

Perhatikan tabel berikut

No	KOH 0,1 M	HI 0,15 M
1	2,00 ml	20,00 ml
2	8,00 ml	20,00 ml
3	15,00 ml	20,00 ml
4	25,00 ml	20,00 ml
5	30,00 ml	20,00 ml

Data di atas merupakan data hasil titrasi dari larutan KOH 0,10 M dengan larutan HI 0,15 M.

Berdasarkan data tersebut maka terjadinya titik netralisasi terletak pada nomor....

- 1
- 2
- 3
- 4
- 5

Gambar 1. Pertanyaan Soal (*frist tier*) Materi Titrasi Asam Basa

Pertanyaan soal (*frist tier*) berbentuk pilihan ganda tradisional yang berisikan konsep dari titrasi asam basa yang diujikan. Soal sudah dilengkapi dengan stimulus soal berupa tabel yang berisikan data hasil titrasi asam basa. Dari data tersebut peserta didik diharapkan dapat menganalisis pertanyaan soal dengan baik.

Alasan yang tepat mengenai pernyataan yang benar di atas adalah...

- Selisih antara titik ekuivalen dengan dengan titik akhir titrasi harus di usahakan seminimal mungkin.
- Titik akhir titrasi adalah titik saat indikator asam-basa mengalami perubahan warna.
- Titik netralisasi adalah ketika mol ekuivalen dari asam kuat yang bereaksi tepat sama dengan nilai mol ekuivalen dari basa kuat.
- Titik ekuivalen dapat diamati dengan melihat perubahan warna pada titrat berdasarkan jenis indikator yang digunakan.
- Titik netralisasi terjadi ketika mol ekuivalen dari asam kuat yang bereaksi lebih banyak dari pada nilai mol ekuivalen dari basa kuat.

Gambar 2. Alasan Soal (*second tier*) Materi Titrasi Asam Basa

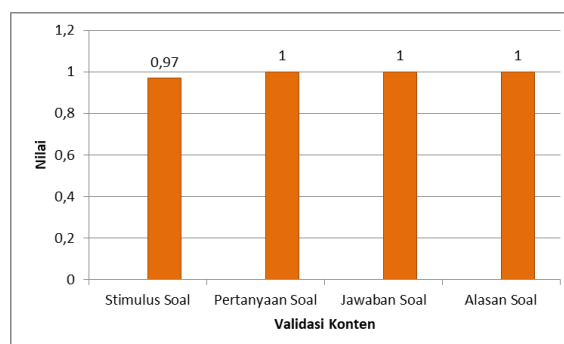
Alasan soal (*second tier*) merupakan alasan jawaban dari pertanyaan soal. Pada

alasan soal (*second tier*) menyerupai format dari soal pilihan ganda tradisional tetapi bertujuan untuk mendorong pemikiran dan penalaran keterampilan berpikir yang lebih tinggi (*Higher Order Thinking Skill*).

Melakukan perbaikan

Melakukan perbaikan untuk memastikan bahwa tes yang akan digunakan dapat berfungsi dengan baik dengan melakukan uji validitas, reliabilitas, daya beda dan indeks kesukaran. Sebuah instrumen tes dikatakan valid apabila tes tersebut dapat mengukur apa yang hendak diukur (Latisma, 2011). Suatu tes dikatakan memiliki validitas apabila tes tersebut tepat, benar, dan shahih atau absah telah dapat menangkap dan mengukur (Sudjono, 2011).

Instrumen tes di validasi oleh validator dengan menyerahkan pernyataan pengetahuan preposisi, peta konsep, dan intrumen tes HOTS kepada validator serta angket yang berisi penilaian terhadap instrumen yang di berikan. Pada validasi isi mempunyai empat bentuk penilaian diantaranya stimulus soal, pertanyaan soal (*frist tier*), jawaban soal, dan alasan soal (*second tier*). Setelah validasi isi instrumen tes selesai, hasil data di analisis dengan menggunakan formula CVR dengan hasil sebagai berikut :

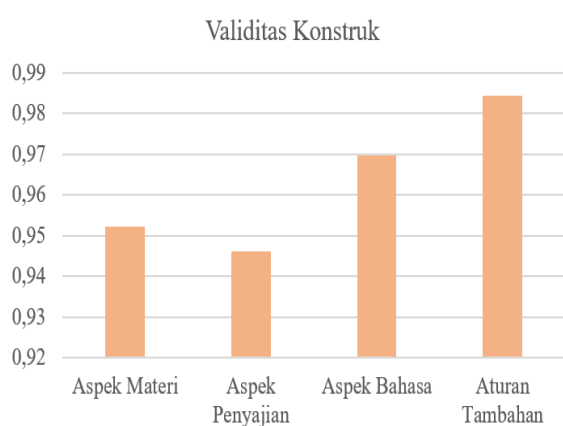


Gambar 3. Analisis Validasi konten Instrumen Tes

Berdasarkan hasil analisis, diperoleh nilai CVR pada stimulus soal yakni 0,97. 2 soal diantaranya belum menggambarkan ciri khas dari soal HOTS. Hal ini disebabkan stimulus yang dipaparkan

belum menunjukkan karakteristik dari soal HOTS. Namun secara keseluruhan soal dapat dikatakan valid.

Pada validitas konstruk, data dianalisis dengan menggunakan *Aiken's V*. penilaian dilakukan dengan 4 aspek yang dinilai yaitu dari aspek materi, penyajian, bahasa, dan tambahan, berdasarkan hasil analisis validitas konstruk yang terdiri atas aspek materi, penyajian, Bahasa, dan aturan tambahan berturut-turut sebesar 0,95, 0,95, 0,97, 0,98.



Gambar 4. Analisis Validasi Konstruk Instrumen Tes

Tahapan selanjutnya yaitu uji coba instrumen tes kepada 20 orang peserta didik. Tahapan ini bertujuan untuk mendapatkan nilai reabilitas, tingkat kesukaran dan daya beda dari instrumen tes HOTS. Analisis butir soal di analisis menggunakan aplikasi Anates. Berdasarkan hasil analisis butir soal menggunakan aplikasi Anates, nilai reabilitas dibagi menjadi dua yaitu reabilitas dari *first tier* (pertanyaan soal) dan *second tier* (alasan soal). Nilai reabilitas dari *first tier* (pertanyaan soal) adalah 0,89 sedangkan untuk *second tier* (alasan soal) adalah 0,90. Suatu instrumen dapat dikatakan reliabel apabila nilai dari reabilitasnya diatas 0,7 (Sudjono, 2011). Sehingga dapat disimpulkan bahwa instrumen tes yang telah dibuat memiliki nilai reabilitas yang tinggi.

Butir-butir soal yang baik adalah butir soal yang tidak terlalu sukar dan tidak terlalu mudah, dengan kata lain memiliki derajat kesukaran yang sedang (Arikunto, 2013). Tingkat kesukaran pada *first tier* (pertanyaan soal) yaitu 18 soal berada pada tingkat sedang, dan 7 soal berada pada tingkat sukar. Hasil analisis tingkat kesukaran soal dari *second tier* (alasan soal) menunjukkan bahwa 2 soal berada pada tingkat sukar, sedangkan 23 soal berada pada tingkat sedang, sehingga dapat disimpulkan bahwa tingkat kesukaran pada instrumen tes tergolong sedang.

Analisis daya beda bertujuan untuk membedakan peserta didik yang memiliki tingkat kemampuan tinggi dengan tingkat kemampuan rendah (Arikunto, 2013). Butir-butir soal yang baik adalah butir-butir soal yang mempunyai indeks diskriminasi 0,4 sampai 0,7. Instrumen tes HOTS yang baik adalah instrumen yang sebagian besar daya pembeda termasuk kategori sedang, baik, dan baik sekali. Instrumen tes yang memiliki daya beda yang sedang, baik, dan baik sekali sudah dapat membedakan antara peserta didik yang berkemampuan tinggi dan rendah (Latisma, 2011).

Tabel 2. Persentase Daya Beda Instrumen Tes

Instrumen Tes	Persentase			
	Jelek	Sedang	Baik	Baik Sekali
<i>First Tier</i> (pertanyaan soal)	8 %	12%	36%	44%
<i>Second Tier</i> (alasan soal)	24%	8%	16%	52%

Berdasarkan hasil analisis daya beda terhadap instrumen tes diatas menunjukkan bahwa persentase daya beda jelek pada *first tier* (pertanyaan soal) hanya 8% sedangkan untuk *second tier* (alasan soal) hanya 24%. Hal ini menunjukkan bahwa instrumen tes *two-tier Higher Order Thinking Skills* berbasis IBT pada materi titrasi asam basa sudah mempunyai daya soal yang efektif.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa instrument tes *two-tier Higher Order Thinking Skills (HOTS)* berbasis IBT pada materi titrasi asam basa yang dikembangkan sudah valid dan reliabel serta memiliki daya beda dan indeks kesukaran yang baik. Sehingga instrumen tes sudah dapat digunakan.

KETERBATASAN DAN IMPLIKASI UNTUK PENELITIAN LAIN

Penelitian ini dibatasi sampai validitas dan praktikalitas. Jumlah sampel yang digunakan yaitu 20 orang siswa yang sampai pada tahap uji coba skala kecil. Oleh karena itu penulis menyarankan agar dilakukannya penelitian dengan uji skala besar menggunakan instrumen tes *two-tier Higher Order Thinking Skill (HOTS)* berbasis IBT pada materi titrasi asam basa.

REFERENSI

- Aiken, L. (1985). *Three Coefficients for Analyzing the Reliability and Validity of Ratings*. 131–142.
- Anderson, L., & Krothwhl, D. (2010). *Kerangka Landasan Untuk Pembelajaran, Pengajaran, dan Assesment, Revisi Taksonomi Pendidikan Bloom. (Terjemahan Agung Prihantoro)*. Addison Wesley Longman, Inc.
- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Kemendikbud. (2017). *Sistem Penilaian Hasil Belajar dan Kemampuan Guru Melaksanakan Penilaian Berdasarkan Kurikulum 2013*. Pusat Penelitian Kebijakan Pendidikan dan Kebudayaan, Balitbang.
- Latisma. (2011). *Evaluasi Pendidikan*. UNP Press.
- Lawshe. (1985). A Quantitative Approach to Content Validity. *Personel Psyclogy*.
- Maulidiansyah, D. (2018). Pengembangan Tes Diagnostik Menggunakan Aplikasi Google Form Materi Momentum dan Impuls Untuk Siswa SMA. *Jurnal Pendidikan Dan Pembelajaran Khatulistiwa*, 7(7), 1–9.
- Nugroho, R. A. (2021). *HOTS (Kemampuan Berpikir Tingkat Tinggi) Edisi Revisi*. PT Gramedia Widiasarana Indonesia.
- Partinem. (2019). *Modul Penyusunan Soal Keterampilan Berpikir Tingkat Tinggi (Higher Order Thinking Skills)* (L. Cahya Wibawa Lusiana (ed.)). Direktorat Pembinaan Sekolah Menengah Atas Direktorat Jenderal Pendidikan Dasar dan Menengah Kementerian Pendidikan dan Kebudayaan.
- Purwanto. (2011). *Evaluasi Hasil Belajar*. Pustaka Pelajar.
- Sudjono, A. (2011). *Pengantar Evaluasi pendidikan*. Rajawal.
- Thomas, A., & Thorne. (2009). *How to increase to higher order thinking*.
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159–169. <https://doi.org/10.1080/0950069880100204>
- Tuysuz, C. (2009). Development of two-tier diagnostic instrument and assess students' understanding in chemistry. *Scientific Research and Essays*, 4(6), 626–631.
- Wilson, F., Pan, W., & Schumsky, D. A. (2012). *Measurement and Evaluation in Counseling and Development*. <https://doi.org/10.1177/074817561244>

0286