

ORBITAL : JURNAL PENDIDIKAN KIMIA

Website : jurnal.radenfatah.ac.id/index.php/orbital

ISSN 2580-1856 (print) ISSN 2598-0858 (online)

PENGEMBANGAN *THREE TIER MULTIPLE CHOICE TEST* PADA MATERI KESETIMBANGAN KIMIA MATA KULIAH KIMIA DASAR LANJUT

Pandu J Laksono

Program Studi Pendidikan Kimia Universitas Islam Negeri Raden Fatah Palembang

E-mail: pandujati_uin@radenfatah.ac.id

ARTICLE INFO

Article History:

Received April 2020

Revised form June 2020

Accepted June 2020

Published online June 2020

Abstract: This study aims to develop a three tier multiple choice test instrument on chemical equilibrium material to identify misconceptions in students. The development model used by Borg and Gall (1983): (1) research and data collection, (2) planning, (3) developing product drafts, (4) conducting initial field trials, (5) making revisions to initial products; (6) conducting limited field tests, (7) product improvement. This research was conducted with 60 respondents. Validation uses the Aiken method with 5 expert validators. The conclusion is that the three tier multiple choice test instrument developed was declared to be feasible and met the criteria as a good problem with the average aiken validity of 0.87. The instrument was declared feasible in terms of the reliability of the test 0.806 included in the high category, has a distinguishing power of 0.351 so that it was categorized as good. Difficulty level found 20% categorized as easy problems, 71.11% categorized as moderate and 8.89% categorized as difficult. Based on the deception index, it was concluded that the deception mostly worked better than the answer key and the deception value mostly more than 5% was chosen so that it was declared effective. Three Tier Multiple Choice Test instrument seen from the practicality included in good categories with a percentage of 78.28%.

Keywords: advanced basic chemistry, chemical equilibrium
three tier multiple choice test

Abstrak: Penelitian ini bertujuan untuk mengembangkan instrumen *Three Tier Multiple Choice Test* pada materi kesetimbangan kimia untuk mengidentifikasi miskonsepsi pada mahasiswa. Model pengembangan yang dipakai Borg dan Gall (1983):(1) penelitian dan pengumpulan data;(2) melakukan perencanaan;(3) mengembangkan draft produk;(4) melakukan uji coba lapangan awal;(5) melakukan revisi terhadap produk awal;(6) melakukan uji lapangan terbatas;(7) penyempurnaan produk. Penelitian ini dilakukan dengan jumlah responden sebanyak 60. Validasi menggunakan metode Aiken dengan 5 orang validator ahli. Kesimpulan adalah Instrumen *Three Tier Multiple Choice Test* yang dikembangkan dinyatakan layak dan memenuhi kriteria sebagai suatu soal yang baik dengan hasil validitas aiken rata-rata 0,87. Instrumen dinyatakan layak ditinjau dari reliabilitas tes 0,806 termasuk di kategori tinggi, memiliki daya pembeda 0,351 sehingga dikategorikan baik. Tingkat kesukaran didapatkan 20% dikategorikan soal yang mudah, 71,11% dikategorikan sedang dan 8,89% dikategorikan sulit. Berdasarkan indeks pengecoh disimpulkan pengecoh sebagian besar bekerja lebih baik daripada kunci jawaban dan nilai pengecoh sebagian besar lebih dari 5% dipilih sehingga dinyatakan efektif. Instrumen *Three Tier Multiple Choice Test* dilihat dari kepraktisan termasuk dalam kategori baik dengan presentase 78,28%.

Keywords: kesetimbangan kimia, kimia dasar lanjut, *three tier multiple choice test*

PENDAHULUAN

Kimia dasar merupakan salah satu mata kuliah yang wajib diambil oleh mahasiswa prodi pendidikan kimia. Kimia umumnya mempelajari bangun (struktur) materi dan perubahan-perubahan yang dialami materi ini dalam proses-proses ilmiah maupun dalam eksperimen yang direncanakan (Bromberg, Katz, & Silberberg, 2009). Kimia yang dikenal merupakan komposisi zat dan penggunaan bahan-bahan tak bersenyawa, baik alamiah maupun buatan dan mengenal proses-proses penting dalam benda hidup, termasuk tubuh kita sendiri. (Keenan, Kleinfelter, & Wood, 1992). Kean & Middlecamp (1985) perspektif kimiawi dapat dikembangkan lewat pengamatan eksperimen kita sendiri. mengatakan bahwa salah satu karakteristik ilmu kimia adalah bersifat abstrak, berurutan dan berjenjang. Hal ini berarti untuk memahami konsep kimia yang lebih kompleks diperlukan pemahaman yang benar pada konsep dasarnya. Apabila mahasiswa mengalami kesulitan pada salah satu konsep dasar, maka terdapat kemungkinan mahasiswa mengalami kesulitan terhadap konsep yang lebih kompleks. Pada faktanya banyak terdapat materi kimia dalam kimia dasar yang tidak mudah dipahami. Mata kuliah Kimia Dasar bukanlah hal yang asing bagi mahasiswa karena semasa Sekolah Menengah Atas telah mendapatkan cukup materi tentang kimia secara umum.

Kemampuan dasar yang diharapkan dimiliki oleh mahasiswa pada semester satu tentunya dapat menjadi modal yang digunakan untuk mempermudah pencapaian tujuan pembelajaran yang berhubungan pada prestasi belajar. Berdasarkan fakta dilapangan, ditemukan bahwa antar kenyataan dan harapan masih terpaut sangat jauh. Masih banyak mahasiswa baru yang mengambil mata kuliah ini masih memperoleh nilai dibawah standar dan banyak mengalami kesulitan (Faika & Side, 2011). Kesulitan dalam kimia bukan hanya dialami oleh peserta didik di tingkat Sekolah Menengah Atas, tetapi juga dialami pada mahasiswa pendidikan kimia. Mahasiswa pendidikan kimia merupakan calon pendidik sehingga kemampuan memahami konsep dengan sangat baik menjadi prioritas yang tidak terelakkan. Pada Materi mata kuliah kimia dasar tak sedikit mahasiswa yang tidak memahami apa yang disampaikan oleh dosen pengampu (Faika & Side, 2011). Salah satu penyebab awal kesulitan belajar adalah anggapan bahwa kimia merupakan salah satu mata pelajaran yang sulit dipahami. Kebanyakan dari mereka yang mengalami kesulitan juga mengalami kesalahan konsep dalam materi. Pemahaman tidak tepat yang melekat mempengaruhi konsep belajar sehingga pada akhirnya berdampak pada rendahnya hasil belajar pada perkuliahan mahasiswa. Hal yang juga dapat menjadi ukuran untuk mengetahui penyebab kesulitan belajar mahasiswa adalah hambatan-hambatan yang dialami mahasiswa dalam proses pembelajaran. Hambatan yang dimaksud yaitu; pada kompetensi pendukung dalam penguasaan konsep, pada proses pembelajaran, dan pada faktor lingkungan mahasiswa. Kesulitan belajar merupakan faktor penghambat yang menjadi penyebab kesulitan belajar mahasiswa pada mata kuliah kimia dasar (Faika & Side, 2011)

Arifin (2012), kesulitan dalam mempelajari ilmu kimia dapat bersumber pada: (1) Kesulitan dalam memahami istilah dimana kesulitan ini timbul karena kebanyakan pebelajar hanya hafal akan istilah dan tidak memahami dengan benar istilah yang sering digunakan dalam pengajaran kimia; (2) Kesulitan dalam memahami konsep kimia karena kebanyakan konsep-konsep dalam ilmu kimia maupun materi kimia secara keseluruhan merupakan konsep atau materi yang bersifat abstrak dan kompleks, sehingga pebelajar dituntut untuk memahami konsep-konsep tersebut dengan benar dan mendalam; (3) Kesulitan angka, dalam pengajaran kimia kita tidak terlepas dari perhitungan secara matematis, dimana pebelajar dituntut untuk terampil dalam rumusan matematis.

Kesetimbangan kimia merupakan salah satu materi yang terdapat pada kimia dasar lanjut. Kesetimbangan kimia adalah pertukaran materi dan energi pada dasarnya mungkin dalam dua arah pada saat yang bersamaan, dalam reaksi maju dan mundur (Chang, 2004). Akibatnya, ada dalam kondisi konstan tertentu hubungan yang pasti antara reaktan dan produk yang digambarkan sebagai keseimbangan kimia (Barke, Hazari, & Yitbarek, 2009). Kesetimbangan kimia, dianggap sebagai salah satu konsep kimia yang sulit untuk diajarkan, melibatkan tingkat kesalahpahaman yang tinggi dari pebelajar (Barke dkk., 2009). Materi yang sering terjadi miskonsepsi adalah materi kesetimbangan kimia, karena materi ini merupakan materi dasar untuk penerimaan konsep kimia selanjutnya, misalnya termokimia, laju reaksi, kelarutan dan hasil kali kelarutan (Iriyanti, Mulyani, & Ariani, 2012)

Bergquist & Heikkinen (1990) telah merangkum konsepsi miskonsepsi yang dialami pada kesetimbangan kimia diantaranya sebagai berikut:

1. Kebingungan ditemui pada saat mengenai jumlah (mol) dan konsentrasi (mol / l) dengan mencoba menghitung konsentrasi ketika diberikan molaritas; mengungkapkan ketidakpastian kapan harus menggunakan volume dengan asumsi rasio mol stoikiometrik berlaku di antara konsentrasi produk dan reaktan, dengan asumsi jumlah molar sama bahkan ketika satu kelebihan.
2. Kebingungan ditunjukkan tentang penampilan dan lenyapnya bahan dengan mengasumsikan konsentrasi berfluktuasi ketika keseimbangan terjadi: reaksi dapat dibalik namun berjalan sampai selesai, bahwa reaksi maju harus diselesaikan sebelum reaksi sebaliknya dimulai, dan penambahan lebih banyak reaktan hanya mengubah konsentrasi produk.
3. Kebingungan ditunjukkan tentang makna K_c dengan menggambarannya sebagai nilai yang bervariasi sementara pada suhu konstan, dengan asumsi bahwa nilainya berubah dengan jumlah produk reaktan.
4. Kebingungan ditunjukkan tentang penggunaan prinsip Le Chatelier dengan berusaha menyesuaikan sistem yang sudah pada keseimbangan; untuk mengubah konsentrasi reaktan yang ditambahkan saja; untuk mengubah nilai konsentrasi semua spesies yang ada kecuali reaktan yang ditambahkan; ketidakpastian bagaimana perubahan suhu, volume, atau tekanan (termasuk penambahan gas yang tidak bereaksi) akan mengubah konsentrasi kesetimbangan.

Konsepsi dalam kimia didasarkan pada pendekatan konstruktivis untuk belajar, dimana mahasiswa membangun struktur kognitif mereka sendiri. Menurut pendekatan pembelajaran ini, peserta didik menghasilkan makna mereka sendiri berdasarkan latar belakang mereka, sikap, kemampuan, pengalaman, sebelum, selama dan setelah pembelajaran. Pembangunan konsep yang dilakukan sendiri, memungkinkan terjadinya perbedaan konstruksi dari yang mereka pahami dengan yang disajikan oleh pengajar (Barke dkk., 2009). Beberapa kasus miskonsepsi yang ditemui dalam materi ini adalah asumsi bahwa katalis mengarah pada produk hasil yang lebih tinggi yang meningkatkan energi aktivasinya padahal konsep yang benar adalah katalis memperbesar laju reaksi karena akan menurunkan energi pengaktifan. Penurunan energi pengaktifan tersebut akan berlaku untuk kedua arah. Seseorang dapat mengalami berbagai macam kesalahpahaman karena kesulitan dalam mengajarkan mata pelajaran serta untuk memahaminya permasalahan yang dimaksudkan (Barke dkk., 2009)

Kesalahan konsep dari mahasiswa dapat diketahui dengan menggunakan tes soal essay dan pilihan ganda (Murni, 2013). Soal essay pada kenyataannya sering terjadi subjektivitas penilaian sehingga dimungkinkan hasilnya akan bias tes yang dapat digunakan untuk menanggulangi kebiasaan tersebut dapat dengan tes pilihan ganda (Rosyana, Ashadi, & Mulyani, 2019). Kelebihan lain soal objektif tes adalah waktu yang digunakan dalam penelitian ini dirasa relatif singkat dibanding tes dengan soal essay akan tetapi tes jenis ini memiliki kelemahan diantaranya adalah jawaban asal dan *gambling* penjawab soal (Laksono, 2019). Dewasa ini telah banyak berkembang soal pilihan ganda yang dapat mengukur kemampuan maupun mengidentifikasi kesalahan yang dilakukan

oleh peserta didik. Pada evaluasi hasil perkuliahan dosen menunjukkan masih minimnya penggunaan instrumen untuk mengungkap kesalahan konsep atau miskonsepsi yang bisa diketahui dengan tes *multiple choice*. Kelemahan soal *multiple choice* bisa dibenahi sebagai identifikasi kesalahan konsep dengan cara dikembangkan menjadi soal *multiple choice* bertingkat yang dapat digunakan sebagai tes diagnostik (Chandrasegaran, Treagust, & Mocerino, 2007). Tes diagnostik adalah salah satu caranya, dari hal diatas pengertian dari tes diagnostik adalah tes yang bertujuan untuk mengidentifikasi kesulitan belajar siswa dalam hal memahami konsep-konsep kunci pada topik tertentu (Suwanto, 2013). Instrumen tes diagnostik yang valid dan andal dapat digunakan untuk mengukur pemahaman konseptual siswa tentang keadaan dan materi (Kirbulut & Geban, 2014). Identifikasi yang cepat memungkinkan dalam penanganan yang baik dan cepat sehingga kesalahan dalam pemahaman konsep dapat ditanggulangi dengan lebih cepat. Waktu menjadi sangat penting karena keterbatasan waktu dalam perkuliahan sering sekali dialami. Berdasarkan banyaknya masalah diatas maka peneliti beranggapan bahwa perlunya suatu instrumen yang dapat dengan cepat mengidentifikasi kesalahan pada mahasiswa.

Bentuk instrumen yang dapat digunakan salah satunya adalah tes diagnostik. Prinsip dasar tes diagnostik yaitu pengajar harus mempertimbangkan pengetahuan intuitif dasar yang telah peserta didik bangun jika ingin memahami pemikiran peserta didik terkait konsep-konsep ilmu pengetahuan yang telah diajarkan (Mubarak, 2016). Desain tes diagnostik pendeteksi miskonsepsi disusun berdasarkan tingkat proses berpikir (Nurfainzani, Susilaningih, & Jumaeri, 2014). Desain tes diagnostik pendeteksi miskonsepsi dirancang sedemikian rupa untuk dapat digunakan menganalisis pemahaman konsep dasar siswa secara nyata dan langsung dapat diketahui seberapa jauh penguasaan konsep, seberapa dalam penguasaan materi yang telah diberikan selama pembelajaran (Bayrak, 2013). Tes diagnostik digunakan untuk menentukan bagian mana saja pada suatu mata pelajaran yang memiliki kelemahan dan menyediakan alat untuk menemukan penyebab kekurangan tersebut serta digunakan untuk mengetahui kekuatan dan kelemahan peserta didik dalam belajar (Suwanto, 2013).

Tes diagnostik yang baik dapat memberikan gambaran yang akurat tentang miskonsepsi yang dimiliki siswa berdasarkan informasi kesalahan yang dibuatnya. Instrumen diagnostik *two-tier* telah dikembangkan mampu mengidentifikasi miskonsepsi siswa pada materi stoikiometri, hidrokarbon, dan laju reaksi (Siswaningsih, 2014). *Three tier multiple choice test* dapat dikembangkan sesuai dengan instrumen penilaian *two-tier* (Chandrasegaran dkk., 2007). Ada berbagai alasan untuk menggunakan instrumen *three tier multiple choice test*, terlepas dari objektivitas dalam penilaian, pengambilan sampel domain konten yang luas, penilaian mekanik, dan generalisasi. Tes diagnostik *three tier multiple choice test* memiliki keunggulan dalam hal memungkinkan para peneliti untuk memeriksa validitas instrumen dan memperkirakan skor kesalahpahaman. Korelasi antara skor dua-tingkat dan tingkat kepercayaan dan persentase negatif palsu dan positif palsu memberikan bukti validitas tes.

Penelitian ini berfokus pada pengembangan instrumen *three tier multiple choice test* pada materi kesetimbangan kimia mata kuliah kimia dasar lanjut.

Pengembangan dilakukan untuk mengetahui penguasaan konsep mahasiswa pada mengatasi kesalahan konsep materi kesetimbangan kimia. Pada penelitian instrumen tes diagnostik instrumen tes pendeteksi miskonsepsi model *three tier multiple choice* dapat digunakan menganalisis pemahaman konsep dasar kimia mahasiswa (Susilaningih, Kasmui, & Harjito, 2016). Tes tiga tingkat memperkirakan skor kesalahpahaman lebih akurat dibandingkan dengan tes satu tingkat dan dua tingkat karena mereka membedakan kesalahpahaman dari kurangnya pengetahuan (Caleon & Subramaniam, 2010; Yeziarski & Birk, 2006; Peşman & Eryilmaz, 2010). Pernyataan ini diperkuat dengan penelitian Kirbulut & Geban (2014) menunjukkan bahwa tes tiga tingkat memprediksi kesalahpahaman siswa lebih akurat dibandingkan dengan tes pilihan ganda dua tingkat dan konvensional karena tes tiga tingkat mencakup skor dua tingkat dan tingkat kepercayaan. Pengembangan instrumen tes yang efektif menjadi salah satu cara untuk menyelesaikan permasalahan tersebut. Penelitian ini diharapkan dapat memberikan informasi yang diperlukan untuk menjadi instrumen alternatif untuk mengidentifikasi kesalahan konsep bagi calon guru kimia.

METODE PENELITIAN

Penelitian dilaksanakan di untuk mengembangkan instrumen *three tier multiple choice test* kepada mahasiswa yang telah mendapatkan materi kesetimbangan kimia dan menempuh kuliah Kimia Dasar Lanjut. Jenis penelitian yang digunakan adalah *Research and Development (R&D)*. Prosedur penelitian dimodifikasi dari model pengembangan Borg & Gall (1983) dan dilakukan hanya sampai pada tahap ketujuh.

Tahapan penelitian dan pengembangan meliputi (1) penelitian dan pengumpulan data (*research and information collecting*) meliputi studi lapangan dan studi literatur; (2) melakukan perencanaan (*planning*) termasuk mengidentifikasi, menyatakan tujuan, menentukan urutan untuk penelitian dan menguji kelayakan skala kecil; (3) mengembangkan draft produk (*develop preliminary form of product*); (4) melakukan uji coba lapangan awal (*preliminary field testing*) dilakukan untuk mengetahui soal-soal mana yang perlu diubah, diperbaiki, bahkan dibuang sama sekali, serta soal-soal mana yang baik digunakan selanjutnya; (5) melakukan revisi terhadap produk awal (*main product revision*) dilakukan setelah melihat kekurangan dari instrumen yang dikembangkan; (6) melakukan uji lapangan terbatas (*main field testing*) dilakukan dengan sejumlah mahasiswa untuk mengetahui kualitas butir soal yang dikembangkan; (7) penyempurnaan produk hasil uji coba (*operasional product revision*). Subjek penelitian ini terdiri atas 60 mahasiswa yang telah menempuh mata kuliah kimia dasar dan mendapatkan materi kesetimbangan kimia.

Data dalam penelitian ini terdiri atas data kualitatif dan kuantitatif. Data kualitatif pada penelitian ini diperoleh dari wawancara dengan dosen pengampu, analisis kebutuhan dari mahasiswa, masukan para validator mengenai instrumen tes yang dikembangkan. Data kuantitatif digunakan untuk mengetahui kualitas instrumen yang dikembangkan ditinjau dari aspek kevalidan, kepraktisan, dan keefektifan. Data kuantitatif pada penelitian ini diperoleh dari hasil penilaian para validator, hasil angket penilaian mahasiswa, hasil angket kepraktisan instrumen,

dan hasil ujicoba instrumen tes secara empirik yang digunakan untuk mengukur reliabilitas, tingkat kesukaran, daya pembeda, dan indeks pengecoh.

Pengujian validitas ahli dilakukan dengan formula aiken. Validasi dilakukan dengan lima raters atau validator. Validator instrumen *three tier multiple choice test* terdiri atas validator ahli bahasa bidang kimia, ahli mata kuliah kimia dasar, ahli materi, ahli evaluasi, dan praktisi. Hasil validasi akan menentukan layak tidaknya sebuah instrumen tes yang dikembangkan sebelum memasuki tahap ujicoba lapangan. Validitas jenis ini merupakan validitas logis dimana sebuah instrumen evaluasi menunjuk pada kondisi yang memenuhi persyaratan valid berdasarkan hasil penalaran (Arikunto, 2011).

Pada ujicoba empirik dengan melakukan analisis (menggunakan program *software ITEMAN*). Rumus yang digunakan adalah rumus *Alpha-Cronbach*:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum s_i^2}{s_t^2} \right)$$

Keterangan :

α = koefisien reliabilitas tes

n = banyaknya butir item

s_t^2 = varian total

$\sum s_i^2$ = jumlah variasi skor tiap-tiap butir soal

(Azwar, 2015)

Daya pembeda dihitung melalui selisih jawaban antara proporsi kelompok tinggi yang menjawab benar dengan kelompok bawah yang menjawab benar. Tingkat kesukaran soal dihitung melalui proporsi jawaban keseluruhan penjawab yang benar pada soal. Purwanto (2009) mengungkapkan bahwa sebuah tes yang dapat dikatakan baik sebagai alat pengukuran jika memenuhi persyaratan kualitas tes, yaitu memiliki validitas, reliabilitas, objektivitas, dan praktibilitas yang baik.

Kepraktisan instrumen evaluasi didapatkan dari hasil penilaian oleh pengguna di program studi. Kepraktisan instrumen adalah kemungkinan suatu instrumen digunakan kembali untuk mengukur tujuan pembelajaran pada masa depan (Purwanto, 2009). Analisis data dilakukan dengan cara menghitung skor yang dicapai dari seluruh aspek yang dinilai kemudian menghitungnya dengan rumus sebagai berikut:

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

Keterangan :

N : Persentase kelayakan aspek

k : Skor hasil pengumpulan data

Nk : Skor maksimal (skor kriteria tertinggi x jumlah aspek x jumlah quoter)

(Purwanto, 2011)

Tabel 1. Interpretasi Skor Penilaian

Presentase Skor	Kategori Kualitas
$86 \% \leq N < 100\%$	Sangat baik
$72 \% \leq N < 85\%$	Baik
$58 \% \leq N < 71\%$	Cukup
$44 \% \leq N < 57\%$	Kurang
$N \leq 44 \%$	Sangat Kurang

(Sudjana, 2009)

Tabel 2. Soal *Three Tier Multiple Choice*

Tingkatan Soal	Bentuk Soal
<i>Tier 1</i>	Tetapan terjadi pada reaksi bolak-balik yang mana laju terbentuknya reaktan sama dengan laju terbentuknya produk untuk reaksi kesetimbangan berikut : $2A(g) + B(g) \rightleftharpoons C(g) + D(g)$ pada suhu tertentu. Adalah 8. Bila pada suhu tetap dan volume diubah menjadi setengah volume asal maka tetapan kesetimbangannya adalah A. 16 B. 8 C. 4 D. 2 E. $\frac{1}{2}$
<i>Tier 2</i>	Alasan jawaban.... A. Apabila volume diubah menjadi setengah maka tetapan kesetimbangan dikalikan 2 B. Apabila suhu tetap maka tidak terjadi perubahan tetapan kesetimbangan C. Apabila volume dirubah menjadi setengah maka tetapan kesetimbangan dibagi dengan 2 D. Apabila suhu tetap dan Volume setengah maka tetapan kesetimbangan juga setengah dari mula-mula E. Tetapan kesetimbangan sama dengan setengah volume asal dibagi dua
<i>Tier 3</i>	Apakah anda yakin dengan jawaban anda : A. Ya B. Tidak

HASIL DAN PEMBAHASAN

A. Hasil Penelitian

1. Studi Lapangan

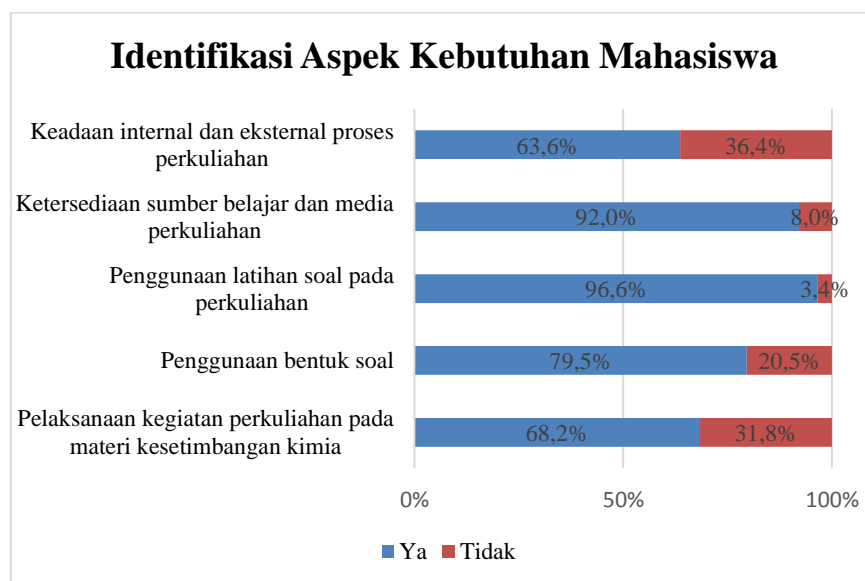
a. Analisis Kebutuhan Dosen

Pemahaman konsep mahasiswa menurut dosen pengampu mata kuliah masih banyak yang mengalami tidak paham konsep dan miskonsepsi pada materi kimia. Tidak paham konsep dan miskonsepsi merupakan kelemahan dalam memahami maka terjadi kesulitan belajar. Pemahaman yang rendah terhadap konsep mengindikasikan adanya kesulitan dalam proses belajar. Kesulitan dalam pembelajaran dapat diketahui dari proses evaluasi. Pada proses ini umumnya dosen melakukan evaluasi pengukuran dengan soal jenis essay. Dosen pengampu mata kuliah pada umumnya lebih sering menggunakan soal essay, dalam

evaluasi karena soal jenis ini lebih dapat mengungkap pengetahuan yang dimiliki siswa secara lebih mendalam dibandingkan dengan soal pilihan ganda, soal jenis ini dianggap dapat mengurangi tindakan curang dari mahasiswa dan sesuai karena mengukur kedalaman pemahaman konsep mahasiswa. Dosen jarang juga menggunakan soal jenis pilihan ganda untuk mengungkap pengetahuan mahasiswanya karena level pengetahuan untuk mahasiswa seharusnya sudah masuk dalam taraf berpikir kritis. Pada evaluasi hasil perkuliahan dosen menunjukkan masih minimnya penggunaan instrumen untuk mengungkap kesalahan konsep atau miskonsepsi yang bisa diketahui dengan tes *multiple choice* akan tetapi memiliki kelemahan seperti diatas. Soal *multiple choice* dapat digunakan sebagai identifikasi kesalahan konsep. Tes diagnostik adalah salah satu caranya, dari hal diatas pengertian dari tes diagnostik adalah tes yang bertujuan untuk mengidentifikasi kesulitan belajar siswa dalam hal memahami konsep-konsep kunci pada topik tertentu (Suwanto, 2013).

b. Analisis Kebutuhan Mahasiswa

Pada tahapan analisis kebutuhan dilakukan dengan penyebaran angket jawaban disertai alasan. Pada persebaran angket tersebut didapatkan hasil sebagai berikut :



Gambar 1. Identifikasi Aspek Kebutuhan Mahasiswa

Identifikasi aspek kebutuhan mahasiswa diperlukan untuk mengetahui sejauh mana instrumen yang akan dikembangkan dibutuhkan bagi mahasiswa. Aspek tersebut: 1) keadaan internal dan eksternal proses perkuliahan secara persentase 63,6% mengalami masalah dalam perkuliahan baik faktor dari dalam maupun luar. Faktor internal seperti intelegensi, perhatian, minat dan bakat, motif dan motivasi. Faktor eksternal pengajar, metode pembelajaran, sarana dan fasilitas; 2) ketersediaan sumber belajar dan media perkuliahan secara umum 92%

mahasiswa merasa sudah tersedia dengan baik dan layak sehingga penggunaan media sumber belajar dan media dirasa sudah baik; 3) penggunaan latihan soal dalam perkuliahan mendapatkan persentase yang mencapai 96,6%, mahasiswa merasa banyak latihan soal yang telah diberikan dosen dengan berbagai macam tipe soal dari uraian dan pilihan ganda; 4) penggunaan bentuk soal 79,5% mahasiswa menyatakan bentuk soal berbeda akan tetapi belum ada jenis soal pilihan ganda bertingkat atau soal uraian dengan alasan pilihan jawaban; 5) pelaksanaan kegiatan perkuliahan pada materi kesetimbangan kimia dinyatakan efektif sebesar 68,2% berdasarkan angket yang dibagikan kepada mahasiswa.

Pada tahapan analisis kebutuhan mahasiswa didapatkan beberapa hal penting yang menjadi kebutuhan: 1) materi yang sering dianggap sulit diantaranya: kesetimbangan, kinetika kimia, asam dan basa. Menurut mereka materi ini konsepnya cukup mendalam dan dipadukan dengan perhitungan matematis sehingga mahasiswa terkadang merasa kurang memahami makna dari tiap materi yang disajikan, 2) soal essay punya kelemahan subjektifitas dan lama dalam proses pengkoreksian padahal dalam mengidentifikasi konsepsi dapat dilakukan dengan cepat bila digunakan soal *multiple choice*, 3) mahasiswa merasa perlu adanya pengembangan soal karena soal yang diberikan umumnya oleh dosen adalah essay. Beberapa mahasiswa menyatakan soal pilihan ganda akan lebih cepat dalam pengerjaannya, 4) mahasiswa merasa paham akan tetapi saat ditanya tentang materi beberapa tidak paham dan salah paham tentang konsepsi kimia.

2. Studi Literatur

Pemahaman merupakan hasil dari kegiatan berpikir dari apa yang telah diterima. Seorang dengan pemahaman yang baik akan mampu memecahkan masalah dengan baik, ketika seseorang bertindak berdasarkan ingatan tanpa memahami, maka dia dapat melakukan kesalahan (Olson & Hergenhahn, 2015). Belajar berdasarkan pemahaman akan lebih dalam dan dapat digeneralisasikan daripada belajar hanya berdasarkan ingatan tanpa adanya pemahaman. Konsep dan materi pada kesetimbangan kimia diantaranya pengertian kesetimbangan, konstanta kesetimbangan (K_c), konstanta kesetimbangan (K_p), kesetimbangan homogen dan heterogen, faktor-faktor yang mengubah pergeseran kesetimbangan: prinsip le châtelier, pengaruh katalis pada kesetimbangan, hubungan laju reaksi dengan kesetimbangan.

Miskonsepsi adalah konsepsi siswa yang tidak cocok dengan konsepsi para ilmuwan (Suwanto, 2013). Miskonsepsi dapat merupakan pengertian yang tidak akurat tentang konsep, penggunaan konsep yang salah, klasifikasi contoh-contoh yang salah tentang penerapan konsep, pemaknaan konsep yang berbeda, kekacauan konsep konsep yang berbeda dan hubungan hirarkis konsep-konsep yang tidak benar. Wafiyah (2012) mengatakan bahwa miskonsepsi diartikan sebagai konsepsi siswa yang tidak cocok dengan konsepsi para ilmuwan, hanya dapat diterima pada kasus-kasus tertentu dan tidak berlaku untuk kasus-kasus lainnya serta tidak dapat digeneralisasikan.

Tes diagnostik adalah salah satu caranya, dari hal diatas pengertian dari tes diagnostik adalah tes yang bertujuan untuk mengidentifikasi kesulitan belajar siswa dalam hal memahami konsep-konsep kunci pada topik tertentu (Suwarto, 2013). Diagnosis kesulitan belajar siswa lebih luas dari pada pelaksanaan tes diagnostik, sehingga dalam pelaksanaan diagnosis kesulitan belajar, selain pelaksanaan tes, perlu dilakukan kegiatan lain, yaitu penelusuran jenis, sumber serta penyebab kesalahan. Tes diagnostik yang baik dapat memberikan gambaran yang akurat tentang miskonsepsi yang dimiliki siswa berdasarkan informasi kesalahan yang dibuatnya. Penelitian yang dilakukan oleh Abbas instrumen *three tier diagnostic test* miskonsepsi suhu dan kalor, mampu mengidentifikasi miskonsepsi siswa pada materi suhu dan kalor, mampu menunjukkan letak kesalahan siswa, membantu guru dalam melaksanakan tes diagnostik pada materi suhu dan kalor, serta membantu siswa dalam memantapkan konsep suhu dan kalor yang dimiliki siswa (Abbas, 2016).

3. Validasi Produk

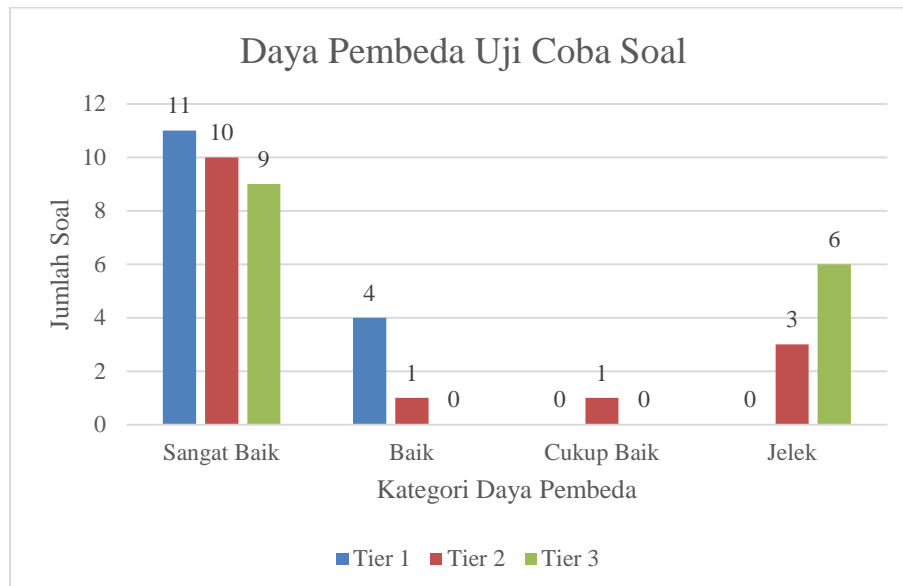
Tujuan utama dari penelitian pengembangan adalah untuk menghasilkan instrumen *three tier multiple choice test* yang telah direvisi berdasarkan masukan para ahli dan data yang diperoleh dari uji coba. Kegiatan pada tahap ini adalah penilaian para ahli (validasi), uji coba, dianalisis kemudian direvisi sesuai dengan tingkat kesukaran, dan daya pembeda. Instrumen penilaian harus memiliki kualitas yang baik. Instrumen penilaian dapat dikatakan memiliki kualitas yang baik yaitu memenuhi penilaian memenuhi persyaratan tersebut dapat dilakukan dua cara, yaitu analisis soal secara teoritik atau kualitatif dan analisis soal secara empiris atau kuantitatif (Sukiman, 2012). Nilai V aiken yang dijadikan standar layak tidaknya suatu instrumen dengan menggunakan 5 validator dan 4 skala pengukuran berdasarkan V tabel Aiken adalah 0,87. Sebanyak 15 butir soal yang diuji validitasnya menggunakan formula Aiken menunjukkan hasil rata-rata lebih dari 0,87. artinya semua soal yang dikembangkan dapat dikatakan Valid secara isi. Validnya instrumen yang dikembangkan secara isi ini merupakan salah satu indikator kualitas instrumen yang dikembangkan. Dari hasil validasi diperoleh hasil 15 soal yang divalidasi dinyatakan layak dan soal tersebut perlu diperbaiki berdasarkan saran dan masukan validator. Selain validitas secara logis juga dilakukan validasi secara empirik yaitu dengan menguji empiris dilakukan dengan cara uji coba untuk mengetahui reliabilitas, daya pembeda, tingkat kesukaran, dan indeks pengecoh soal.

Reliabilitas dihitung pada masing-masing *tier* (tingkat). Tabel 3 menunjukkan bahwa reliabilitas instrumen pada ujicoba soal pada tiap *tier* memiliki perbedaan nilai. Dapat dilihat bahwa pada pengembangan instrumen *three tier multiple choice test* untuk tingkat ke-1, tingkat ke-2, dan tingkat ke-3 nilai reliabilitasnya antara 0,71 – 0,90 dan dikategorikan mempunyai reliabilitas tinggi.

Tabel 3. Reliabilitas Instrumen *Three Tier Multiple Choice Test*

Tier	Uji Coba Soal	
	Reliabilitas	Ket
1	0,721	Tinggi
2	0,804	Tinggi
3	0,884	Tinggi

Daya pembeda adalah kemampuan suatu butir soal sejauh mana dapat membedakan penjawab yang telah menguasai materi dengan penjawab yang belum atau kurang menguasai materi. Daya pembeda ditunjukkan oleh nilai *Point Biser* pada ITEMAN yang merupakan korelasi poin biserial antara jawaban benar per butir soal dengan total skor. Perhitungannya didasarkan pada proporsi kelompok atas yang merupakan kelompok dengan skor tinggi dengan kelompok bawah yang mempunyai skor rendah.

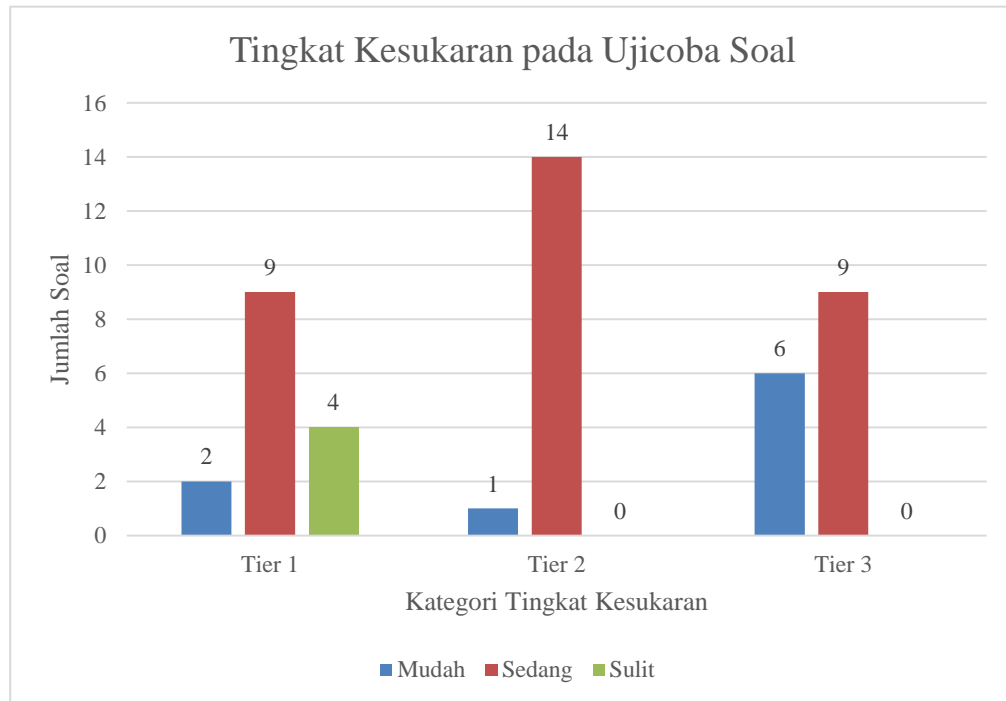


Gambar 2. Daya Pembeda pada Uji Coba Soal

Daya pembeda pada uji coba soal di *tier 1*, *tier 2*, dan *tier 3* sebagian besar sudah termasuk sangat baik. Proporsi soal dengan daya beda cenderung menurun pada *tier 1* sampai *tier 3*. Proporsi soal dengan daya beda jelek juga ditemui pada beberapa soal terutama pada *tier 2* dan *tier 3*. Secara keseluruhan instrumen penilaian *three tier multiple choice test* ini tidak memiliki daya pembeda dengan nilai negatif, artinya instrumen ini memiliki kemampuan untuk membedakan dan lebih banyak mengandung soal yang mampu membedakan kelompok atas dan bawah.

Hasil dari tingkat kesukaran grafik pada gambar 3 menunjukkan bahwa tingkat kesukaran soal pada *tier 1*, *tier 2* dan *tier 3* pada ujicoba soal dengan tingkat kesukaran sedang lebih banyak dibandingkan dengan soal yang sukar dan mudah. Grafik juga menunjukkan bahwa tingkat kesukaran soal pada *tier 2*

dan *tier 3* bahwa tidak ada soal yang dinyatakan sukar. Secara keseluruhan bisa dikatakan bahwa tingkat kesukarannya cenderung sama terutama pada soal kategori sedang merupakan sebagian besar didapatkan dari hasil analisis.

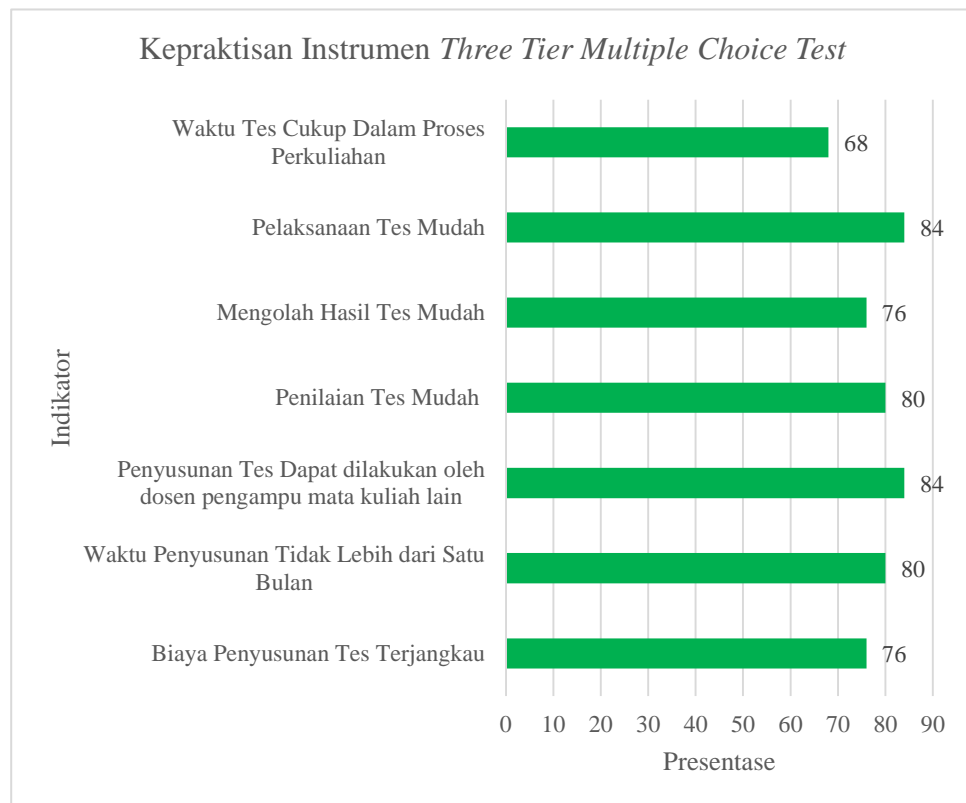


Gambar 3. Tingkat Kesukaran pada Ujicoba Soal

Dalam soal berbentuk pilihan ganda, terdapat pilihan jawaban dimana salah satunya merupakan jawaban yang benar dan yang lainnya salah. Satu jawaban yang benar ini disebut kunci jawaban, sedangkan jawaban lainnya yang salah disebut pengecoh atau distraktor. Pengecoh dikatakan berfungsi dengan baik apabila banyak dipilih oleh siswa, sebaliknya, apabila sedikit yang memilih pengecoh atau sama sekali tidak dipilih maka pengecoh tersebut dikatakan jelek. Pengecoh berfungsi efektif apabila minimal dipilih oleh 5% siswa (Sukiman, 2012). Pada ujicoba kedua ini selain melakukan analisis butir soal yang meliputi reliabilitas, daya pembeda, tingkat kesukaran soal, dan indeks pengecoh dari instrumen penilaian *three tier multiple choice test*, juga dilakukan sebaran angket mengenai penggunaan instrumen dan kepraktisan instrumen *three tier multiple choice test* yang dikembangkan

4. Kepraktisan Produk

Purwanto (2009) mengungkapkan bahwa sebuah tes yang dapat dikatakan baik sebagai alat pengukuran jika memenuhi persyaratan kualitas tes, yaitu memiliki validitas, reliabilitas, objektivitas, dan praktibilitas yang baik. Kepraktisan instrumen evaluasi didapatkan dari hasil penilaian oleh pengguna di program studi



Gambar 4. Kepraktisan Instrumen *Three Tier Multiple Choice Test*

Kepraktisan instrumen *three tier multiple choice test* dilihat dari berbagai indikator: 1) waktu tes yang cukup dalam proses perkuliahan, 2) pelaksanaan tes mudah, 3) mengolah hasil tes mudah, 4) penilaian tes mudah, 5) penyusunan tes dapat dilakukan oleh dosen pengampu mata kuliah lain, 6) waktu penyusunan tidak boleh lebih dari satu bulan, 7) biaya penyusunan tes terjangkau. Berdasarkan dari persentase rata-rata penilaian indikator kepraktisan instrumen evaluasi adalah 78,28% sehingga dapat disimpulkan bahwa kepraktisan termasuk kategori baik. Kepraktisan instrumen evaluasi adalah kemungkinan suatu instrumen evaluasi digunakan kembali oleh pengajar untuk mengukur tujuan pembelajaran pada suatu saat nanti (Purwanto, 2009).

B. Pembahasan

Evaluasi pengetahuan yang umum digunakan dosen di prodi pendidikan kimia adalah dengan tes tertulis. Bentuk soal yang dipakai dalam kuis ataupun ujian semester umumnya adalah soal essay. Soal essay dianggap sebagai soal yang dapat mengukur kemampuan mahasiswa secara lebih mendalam. Soal jenis ini dianggap dapat mengurangi tindakan curang dari mahasiswa. Tes jenis ini memungkinkan mahasiswa menjawab secara bebas. Kebebasan menjawab semacam itu, penskoran tes ini akan sangat subjektif. Pemberian skor dan penilaian sangat tergantung pada dosen pengampu yang memberikan skor tes tersebut. Jawaban terhadap tes ini tidak ditetapkan secara pasti terlebih dahulu

pada saat menyusun pertanyaannya sehingga penskorannya menjadi tidak objektif, maka tes seperti ini disebut tes non-objektif. Ciri khas dari pertanyaan essay adalah kebebasan mahasiswa untuk merespon. Mahasiswa dapat memilih, menghubungkan, dan menampilkan idenya dengan bahasanya sendiri. Disamping meningkatkan nilai kebebasan dari soal essay sebagai alat penilaian prestasi yang kompleks, soal essay juga menyebabkan kesulitan menilai, dan ketidak efektifan dalam menilai kemampuan faktual.

Tes objektif umumnya jarang dipakai di universitas untuk menilai kemampuan mahasiswa karena dirasa rawan akan kecurangan dan asal dalam menjawab dalam pemilihan jawaban. Tes objektif memiliki kelebihan yaitu : 1) dapat digunakan untuk mengukur segala level tujuan pembelajaran, 2) hanya menuntut waktu mengerjakan sangat minimal, maka dapat menggunakan jumlah butir soal yang banyak sehingga pokok bahasan yang diujikan dapat lebih luas, 3) Penskoran hasil tes dapat dilakukan secara objektif (Arikunto, 2007). Beberapa kelebihan ini ingin dimanfaatkan dan ditingkatkan agar dapat digunakan sebagai identifikasi pada pemahaman mahasiswa pada materi yang memungkinkan terjadinya miskonsepsi.

Miskonsepsi pada mahasiswa dapat diketahui dengan tes diagnostik. Tes diagnostik adalah tes yang bertujuan untuk mengidentifikasi kesulitan belajar siswa dalam hal memahami konsep-konsep kunci pada topik tertentu (Suwanto, 2013). Bentuk soal pilihan ganda sekarang bervariasi bentuknya. Salah satu bentuk pilihan ganda yang sering dijumpai atau dikembangkan adalah pilihan ganda beralasan. *Two-tier* adalah salah satu bentuk pilihan ganda beralasan dengan pilihan ganda dua tingkat. Tingkat pertama menyatakan soal dan tingkat kedua alasan menjawab. Alasan dapat berupa jawaban essay atau pilihan ganda. Adapun *two-tier test* dinyatakan sebagai berikut: tingkat pertama merupakan soal yang berkaitan dengan pengetahuan yang menyerupai soal pilihan ganda yang telah ada sebelumnya. Pada tingkat kedua merupakan pilihan alasan-alasan yang mendasari jawaban di tingkat pertama. Persyaratan memilih alasan yang mendasari jawaban di tingkat pertama ini efektif untuk mengukur pembelajaran telah ada atau dikembangkan pilihan ganda tingkat tiga yang disebut *three tier multiple choice test* yang merupakan pengembangan dari tes pilihan ganda. TTMCT ini juga merupakan pengembangan dari *two-tier multiple choice* yang terdiri dari dua tingkat pilihan ganda yang saling berkaitan. Waktu yang minimal dalam pengerjaan tes diagnostik *three tier multiple choice test* menjadi pilihan yang harus dimanfaatkan karena terbatasnya waktu perkuliahan

Pengembangan instrumen yang bagus harus mengalami proses validasi dan dinyatakan valid. Instrumen tes yang dikembangkan harus masuk dalam kategori layak dan memiliki reliabilitas yang dikategorikan baik (Mutammimah, Jumadi, Wilujeng, & Kuswanto, 2019). Penelitian ini merupakan pengembangan tes tingkat tiga di universitas, dimana sebelumnya telah dikembangkan tes serupa dengan tingkat sekolah. Penelitian yang dilakukan Cetin-Dindar & Geban (2011) telah mengembangkan tes diagnostik tiga tingkat yang valid dan dapat diandalkan untuk menilai pemahaman siswa tentang konsep asam-basa menjadi acuan bagi pengembangan instrumen tes yang serupa.

Tes tiga tingkat memprediksi kesalahpahaman siswa lebih akurat

dibandingkan dengan tes pilihan ganda dua tingkat dan konvensional karena tes tiga tingkat mencakup skor dua tingkat dan tingkat kepercayaan. (Kirbulut & Geban, 2014). Instrumen tes diagnostik yang dihasilkan terdiri atas kisi-kisi tes, petunjuk pengerjaan soal, soal tes, kunci jawaban, pedoman penskoran dan pedoman interpretasi hasil. Soal tes terdiri atas tiga tingkatan, yaitu: pertanyaan dengan satu kunci jawaban dan empat pengecoh, pilihan alasan, dan tingkat keyakinan pada pilihan jawaban dan alasan yang diberikan (Mubarak, 2016).

Keunggulan yang dimiliki tes diagnostik *three tier multiple choice* adalah dapat: (1) mendiagnosis miskonsepsi yang dialami peserta didik lebih mendalam, (2) menentukan bagian-bagian materi yang memerlukan penekanan lebih saat pembelajaran, (3) merencanakan pembelajaran yang lebih baik untuk membantu mengurangi miskonsepsi peserta didik (Mubarak, 2016). Pengembangan instrumen *three tier multiple choice test* dapat dikembangkan pada berbagai materi yang terdapat pada perkuliahan kimia. Salah satu materi yang sering terjadi miskonsepsi adalah kesetimbangan kimia (*Chemical Equilibrium*) yaitu pertukaran materi dan energi dalam dua arah pada saat yang bersamaan, dalam reaksi maju dan mundur.

Pengembangan *three tier multiple choice test* pada materi kesetimbangan kimia pada mata kuliah kimia dasar lanjut telah berhasil dikembangkan. Sebagian besar penelitian dan pengembangan instrumen tes berupa tes bertingkat seperti ini digunakan sebagai tes diagnostik terjadinya miskonsepsi. Sejalan dengan penelitian yang telah dilakukan, instrumen jenis *three tier multiple choice* yang dipergunakan untuk mengidentifikasi atau mengungkap miskonsepsi mahasiswa pada konsep optik telah berhasil dikembangkan dengan valid dan reliabel (Pardede, 2016). Pengembangan tes diagnostik *three tier* untuk mengidentifikasi miskonsepsi siswa pada materi ikatan kimia dengan menggunakan program PHP-MySQL berhasil dikembangkan dengan kategori sangat layak (S & Sanjaya, 2015). Dalam menganalisis miskonsepsi diperlukan pedoman dalam melihat yang terjadi adalah paham konsep, miskonsepsi atau tidak paham konsep Tabel 4 menunjukkan analisis miskonsepsi:

Tabel 4. Analisis Miskonsepsi

Tingkat pertama	Tingkat kedua	Tingkat ketiga	Kategori
Benar	Benar	Yakin	Paham konsep
Benar	Benar	Tidak Yakin	Tidak Paham Konsep
Benar	Salah	Yakin	Flase Positif (miskonsepsi)
Benar	Salah	Tidak Yakin	Tidak paham konsep
Salah	Benar	Yakin	Flase negatif
Salah	Benar	Tidak Yakin	Tidak paham konsep
Salah	Salah	Yakin	Miskonsepsi
Salah	Salah	Tidak Yakin	Tidak Paham Konsep

(Gurel, Eryilmaz, & McDermott, 2015)

Implementasi penggunaan tes diagnostik menunjukkan bahwa mahasiswa pada materi kesetimbangan kimia paham konsep sebesar 9 dari 60 mahasiswa (15%), 36 dari 60 mahasiswa (60%) miskonsepsi, dan 15 dari 60 mahasiswa (25%) tidak paham konsep dasar dari kesetimbangan kimia. Penyempurnaan

desain tes diagnostik pendeteksi miskonsepsi, dan perluasan cakupan materi dasar kimia perlu dilakukan agar mendapatkan instrumen yang lebih baik (Susilaningih dkk., 2016)

SIMPULAN DAN SARAN

Simpulan

Instrumen *three-tier multiple choice test* yang dikembangkan dinyatakan layak dan memenuhi kriteria sebagai suatu soal yang baik dengan hasil penilaian dari kelima validator dengan didapatkan validitas aiken dengan nilai rata-rata 0,87. Pada uji empiris *three-tier multiple choice test* pada materi kesetimbangan kimia mata kuliah kimia dasar lanjut yang dikembangkan memiliki reliabilitas tes rata-rata 0,806 dari ketiga *tingkatan soal* yang dikategori tinggi. Daya pembeda rata-rata dari tingkatan soal yaitu *tier 1*, *tier 2*, dan *tier 3* didapatkan rata-rata daya pembeda 0,351 sehingga dapat dikategorikan baik. Tingkat kesukaran didapatkan hasil sebagai berikut : 20% dikategorikan soal yang mudah, 71,11% dikategorikan sedang dan 8,89% dikategorikan sulit. Indeks pengecoh sebagian besar bekerja lebih baik daripada kunci jawaban dan nilai pengecoh sebagian besar lebih dari 5% dipilih sehingga dinyatakan efektif. Kepraktisan instrumen *three tier multiple choice test* setelah dilakukan uji kepraktisan termasuk dalam kategori baik dengan presentase rata-rata 78,28%.

Saran

Saran dalam penelitian ini adalah dalam proses perkuliahan dosen diharapkan melakukan identifikasi atau diagnostik konsep pada materi-materi yang rawan mengalami miskonsepsi. Materi yang sering terjadi miskonsepsi di kimia dasar lanjut adalah kesetimbangan kimia, reaksi redoks, asam basa, ikatan kimia, reaksi senyawa kompleks. Diagnostik menjadi sangat penting karena mahasiswa program studi pendidikan kimia merupakan para calon pendidik di sekolah, sehingga pembelajaran kimia yang memiliki kedalaman konsep dapat dijelaskan tanpa adanya kesulitan kegagalan konsep atau miskonsepsi. Pada penelitian yang lain diharapkan dilakukan juga pada materi kimia dasar lain dengan waktu yang cukup singkat diperkuliahan sehingga dapat segera mengatasi masalah miskonsepsi sebelum evaluasi menyeluruh melalui evaluasi perkuliahan. Penyempurnaan produk penelitian dapat dilakukan dengan memberikan variasi soal yang lebih beragam pada setiap tingkatan soal.

DAFTAR PUSTAKA

- Abbas, M. L. H. (2016). Pengembangan Instrumen Three Tier Diagnostic Test, *01(02)*, 83–92.
- Arifin, Z. (2012). *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya. <https://doi.org/979-692-956-2>
- Arikunto, S. (2007). *Prosedur Penelitian (edisi revisi)*. Jakarta: Rineka Cipta.
- Arikunto, S. (2011). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Buku Aksara.

- Azwar, S. (2015). *Dasar-dasar Psikometri*. Yogyakarta: Pustaka Pelajar.
- Barke, H.-D., Hazari, A., & Yitbarek, S. (2009). *Misconception in Chemistry* (Vol. 52). Berlin: Springer.
- Bayrak, B. K. (2013). Using Two-Tier Test to Identify Primary Students' Conceptual Understanding and Alternative Conceptions in Acid Base. *Mevlana International Journal of Education*. <https://doi.org/10.13054/mije.13.21.3.2>
- Bergquist, W., & Heikkinen, H. (1990). Student ideas regarding chemical equilibrium: What written test answers do not reveal. *Journal of Chemical Education*, 67(12), 1000. <https://doi.org/10.1021/ed067p1000>
- Borg, W.R. & Gall, M.D. (1983). *Educational Research, An Introduction, Fourth Edition*. New York: Log-man Inc.
- Bromberg, Y., Katz, O., & Silberberg, Y. (2009). Ghost imaging with a single detector. *Physical Review A - Atomic, Molecular, and Optical Physics*. <https://doi.org/10.1103/PhysRevA.79.053840>
- Caleon, I., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education*. <https://doi.org/10.1080/09500690902890130>
- Cetin-Dindar, A., & Geban, O. (2011). Development of a three-tier test to assess high school students' understanding of acids and bases. *Procedia - Social and Behavioral Sciences*, 15, 600–604. <https://doi.org/10.1016/j.sbspro.2011.03.147>
- Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The Development of A Two-Tier Multiple-Choice Diagnostic Instrument for Evaluating Secondary School Students' Ability to Describe and Explain Chemical Reactions Using Multiple Levels of Representation. *Chemistry Education Research and Practice*, 8(3), 293. <https://doi.org/10.1039/b7rp90006f>
- Chang, R. (2004). *Kimia Dasar Konsep-Konsep Inti, Jilid 1 edisi 3*. Jakarta: Erlangga.
- Faika, S., & Side, S. (2011). Analisis Kesulitan Mahasiswa dalam Perkuliahan dan Praktikum Kimia Dasar di Jurusan Kimia FMIPA Universitas Negeri Makassar Analysis. *Jurnal Chemica*, 12, 18–26.
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(5), 989–1008. <https://doi.org/10.12973/eurasia.2015.1369a>
- Iriyanti, N. P., Mulyani, S., & Ariani, S. R. D. (2012). Identifikasi Miskonsepsi Pada Materi Pokok Wujud Zat Siswa Kelas VII Smp Negeri 1 Bawang Tahun Ajaran 2009/2010. *Jurnal Pendidikan Kimia*, 1(1), 8–13.

- Kean, E., & Middlecamp, C. (1985). *A Survival Manual for General Chemistry (Panduan Belajar Kimia Dasar)*. Penerjemah: A. Hadyana Pudjaatmaka. Jakarta: Gramedia.
- Keenan, C. W., Kleinfelter, D. C., & Wood, J. H. (1992). *Ilmu Kimia untuk Universitas. Edisi ke-6, Jilid 2*. Jakarta: Erlangga.
- Kirbulut, Z. D., & Geban, O. (2014). Using three-tier diagnostic test to assess students' misconceptions of states of matter. *Eurasia Journal of Mathematics, Science and Technology Education*, 10(5), 509–521. <https://doi.org/10.12973/eurasia.2014.1128a>
- Laksono, P. J. (2019). Pengembangan Dan Penggunaan Instrumen *Two-Tier Multiple Choice* Pada Materi Termokimia Untuk Mengukur Kemampuan Berpikir Kritis. *Orbital: Jurnal Pendidikan Kimia*, 2(2), 80–92. <https://doi.org/10.19109/ojpk.v2i2.2646>
- Mubarak, S. (2016). Journal of Innovative Science Education Pengembangan Tes Diagnostik Three Tier Multiple Choice. *Journal of Innovative Science Education*, 5(2), 101–110.
- Murni, D. (2013). Identifikasi Miskonsepsi Mahasiswa Pada Konsep Substansi Genetika Menggunakan Certainty of Response Index (CRI). *Prosiding Semirata FMIPA Universitas Lampung*, 1(1), 205–212.
- Mutammimah, Jumadi, Wilujeng, I., & Kuswanto, H. (2019). Developing Learning Instruments using Tracker in Measuring Students' Science Process Skills. *Journal of Physics: Conference Series*, 1233(1), 2–11. <https://doi.org/10.1088/1742-6596/1233/1/012053>
- Nurfainzani, P., Susilaningsih, E., & Jumaeri. (2014). Pengembangan Tes Diagnostik Two-Tier Multiple Choice Untuk Mengidentifikasi Miskonsepsi Siswa Kelas Xi. *Chemistry in Education*, 3(2252), 1–7. [https://doi.org/10.1016/S0378-3812\(97\)00311-7](https://doi.org/10.1016/S0378-3812(97)00311-7)
- Olson, M. H., & Hergenhahn, B. R. (2015). *Introduction to theories of learning: Ninth edition. Introduction to Theories of Learning: Ninth Edition*.
- Pardede, H. (2016). Pengembangan *Three Tier Test* Sebagai Instrumen Untuk Mengungkap Miskonsepsi Mahasiswa Pada Konsep Optik. *JURNAL Suluh Pendidikan FKIP-UHN*, 3(2), 148–153.
- Peşman, H., & Eryilmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *Journal of Educational Research*. <https://doi.org/10.1080/00220670903383002>
- Purwanto, N. (2009). *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. PT Remaja Rosdakarya.
- Purwanto, N. (2011). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.

- Rosyana, W., Ashadi, A., & Mulyani, S. (2019). Pengembangan Instrumen Penilaian *Three-Tier Multiple Choice (Three- Tmc)* Untuk Mengukur Keterampilan Berpikir Kritis Siswa Pada Materi Kimia Larutan Penyangga Kelas XI SMA. *INKUIRI: Jurnal Pendidikan IPA*, 8(1), 48–58. <https://doi.org/10.20961/inkuiri.v8i1.31814>
- S, A. C., & Sanjaya, I. G. M. (2015). The Development Of Three Tier Diagnostic Test To Identify Student Misconception In Chemical Bonding On 10Th Grader. *UNESA Journal of Chemical Education*, 4(3), 456–465.
- Siswaningsih, W. (2014). Pengembangan Tes Diagnostik Two-Tier Untuk Mengidentifikasi Miskonsepsi Pada Materi Kimia Siswa Sma. *Jurnal Pengajaran Matematika Dan Ilmu Pengetahuan Alam*, 19(1), 117. <https://doi.org/10.18269/jpmipa.v19i1.487>
- Sudjana, N. (2009). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT. Remaja Rosdakarya.
- Sukiman. (2012). *Pengembangan Sistem Evaluasi*. Yogyakarta: Insan Mandiri.
- Susilaningsih, E., Kasmui, & Harjito. (2016). Desain Instrumen Tes Diagnostik Pendeteksi Miskonsepsi Untuk Analisis Pemahaman Konsep Kimia Mahasiswa Calon Guru. *Unnes Science Education Journal*, 5(3), 1432–1437. <https://doi.org/10.15294/usej.v5i3.13184>
- Suwarto. (2013). *Pengembangan Tes Diagnostik dalam Pembelajaran*. Yogyakarta: Pustaka Pelajar.
- Wafiyah, N. (2012). Identifikasi Miskonsepsi Siswa Dan Faktor- Faktor Penyebab Pada Materi Permutasi Dan Kombinasi Di SMA Negeri 1 Manyar. *Gamatika*, 2(2).
- Yeziarski, E. J., & Birk, J. P. (2006). Misconceptions about the particulate nature of matter. Using animations to close the gender gap. *Journal of Chemical Education*. <https://doi.org/10.1021/ed083p954>