

PENGEMBANGAN INSTRUMEN TES DIAGNOSTIK *FOUR-TIER MULTIPLE CHOICE* PADA MATERI MINYAK BUMI

Basri¹, Ardi widhia sabekti², Nina adriani³

Basrisidik99@gmail.com

Program Studi Pendidikan Kimia, Fakultas Keguruan dan Ilmu Pendidikan,
Universitas Maritim Raja Ali Haji

Abstract

This study aims to develop a Four-Tier Multiple Choice diagnostic test instrument with good criteria in terms of validity and reliability. The development model used was Tessmer's formative research type development model which consists of 3 stages, namely the preparation stage (preliminary), self-assessment (self evaluation), and product manufacture (prototyping). The trial subjects used in this study were students of class XI MIA SMAN 2 Tanjungpinang as many as 12 students at the small group stage. The instruments used in this study were test instruments and non-test instruments. This test instrument was validated by 3 experts using the Aikens V formula. The assessment by the three validators resulted in an average validation of the diagnostic test instrument with 20 items of 0.89 with a very valid category. The results of the analysis on the empirical validation found that there were 2 items that were not valid so as to produce a final prototype with a total of 18 items. The reliability of the instrument obtained is 0.94 with a very high reliability interpretation.

Kata kunci: *Four-Tier Multiple Choice, Reliability, Validity*

I. Pendahuluan

Kimia merupakan cabang Ilmu Pengetahuan Alam (IPA) yang mempelajari struktur, komposisi, sifat, perubahan materi serta energi yang menyertai perubahan tersebut. Dalam penerapan kehidupan sehari-hari ilmu kimia diawali dengan pemahaman konsep, prinsip, dan teori kimia yang benar. Kebanyakan peserta didik gagal dalam memahami konsep atau miskonsepsi dalam belajar kimia yang menuntut pemahaman konsep yang baik (Ardiansah, 2016). Menurut Harahap & Novita (2020) miskonsepsi diartikan sebagai pandangan atau ide yang salah mengenai sebuah konsep yang telah diyakini oleh para ahli kebenarannya. Adapun menurut Mahmudah (2013) miskonsepsi diartikan sebagai kebenaran konsep yang diyakini oleh peserta didik yang tidak sesuai dengan kebenaran sesungguhnya. Keterkaitan antara satu konsep dengan konsep yang lain pada materi kimia membuat miskonsepsi penting untuk diidentifikasi, sehingga tidak menyebabkan miskonsepsi berkelanjutan. Salah satu materi yang dipelajari dalam mata pelajaran kimia yaitu minyak bumi.

Berdasarkan hasil wawancara dengan guru kimia SMAN 2 Tanjungpinang, kesalahan konsep pada materi kimia sering terjadi dan bahkan peserta didik tidak tahu konsep. Materi minyak bumi memerlukan pemahaman konsep yang baik dan tergolong kompleks, dimana karakteristiknya ada yang bisa dilihat secara langsung seperti minyak mentah maupun olahannya, yang tidak bisa dilihat secara langsung contohnya proses terbentuknya minyak bumi serta secara simbolis seperti komponen penyusun minyak bumi. Selanjutnya (Djarwo, 2013) melaporkan bahwa sub bab yang terjadi miskonsepsi peserta didik pada materi minyak bumi diantaranya konsep rumus struktur, reaksi pembakaran senyawa alkana dan reaksi perengkahan. Penggunaan instrumen tes diagnostik mampu mengidentifikasi miskonsepsi pada konsep tertentu (Suparno, 2013).

Menurut Suwanto (2013) tes diagnostik adalah tes yang digunakan untuk mengidentifikasi miskonsepsi atau kelemahan peserta didik sehingga didapat hasil berupa respon untuk diperbaiki kelemahannya. Salah satu jenis instrumen tes diagnostik adalah *four-tier test* atau tes diagnostik empat tingkat yang merupakan pengembangan dari *three tier test* atau tes diagnostik tiga tingkat (Caleon & Subramaniam, 2010). Format instrumen tes pilihan ganda bertingkat empat ini berisikan *tier-1* berupa pokok soal dan pilihan jawaban, *tier-2* berupa keyakinan jawaban, *tier-3* berupa pilihan alasan, dan *tier-4* berupa keyakinan terhadap alasan. Keunggulan *four-tier test* ini mampu mendeteksi tingkat keyakinan peserta didik yang berbeda dalam memilih pilihan jawaban maupun pilihan alasan serta memberikan gambaran informasi yang lebih banyak dibandingkan dengan *three tier test*.

Tujuan pada penelitian ini adalah untuk mengetahui pengembangan instrumen tes diagnostik *four-tier multiple choice* pada materi minyak bumi serta untuk menentukan kelayakan instrumen tes diagnostik *four-tier multiple choice* berdasarkan aspek validitas dan reliabilitas.

II. Metode Penelitian

Jenis penelitian yang digunakan adalah penelitian dan pengembangan (Research and Development / R&D) dengan model pengembangan tipe formative research Tessmer (1993) yang dibatasi sampai pada uji reliabilitas pada kelompok kecil. Tahap pengembangan produk terdiri atas : tahap persiapan (preliminary) yaitu menganalisis kebutuhan dan kajian literatur, tahap penilaian sendiri (self evaluation) dilakukan penilaian sendiri terhadap desain instrumen yang dikembangkan untuk menghasilkan prototipe 1, dan tahap pembuatan prototipe (prototyping) yaitu penelaahan instrumen tes, validasi instrumen tes untuk menghasilkan prototipe 2 yang digunakan untuk pengujian instrumen tes pada kelompok kecil. Subjek penelitian yang digunakan saat uji coba kelompok kecil sebanyak 12 orang peserta didik kelas XI MIA SMAN 2 Tanjungpinang yang telah menerima materi minyak bumi sebelumnya.

Instrumen tes diagnostik yang telah dikembangkan harus dianalisis validitasnya. Pada penelitian ini dilakukan validasi isi dan juga validasi empiris. Data hasil analisis validasi isi diperoleh dari lembar validasi yang dinilai oleh 3 validator yang terdiri dari 1 dosen pendidikan kimia UMRAH dan 2 orang guru kimia SMAN 2 Tanjungpinang. Teknik yang digunakan ialah formula aiken v untuk menghitung content validity coefficient berdasarkan hasil penilaian panel ahli terhadap suatu item. Rumus formula aiken yang digunakan adalah sebagai berikut :

$$V = \frac{\Sigma S}{[n(c-1)]}$$
$$S = r - lo$$

Keterangan :

l_0 = angka penilaian validitas yang terendah

C = angka penilaian validitas tertinggi

r = angka yang diberikan oleh validator

Selanjutnya hasil skor analisis validitas selanjutnya diinterpretasikan berdasarkan kriteria skor yang dapat dilihat pada Tabel 1.

Tabel 1. Kriteria validitas isi

Indeks V Aikens	Kategori
$V < 0,4$	Kurang Valid
$0,4 < V < 0,8$	Valid
$V > 0,8$	Sangat Valid

Sumber: (Azwar, 2012)

Selanjutnya dilakukan uji validitas empiris pada tahap uji coba kelompok kecil. jika $r_{\text{Hitung}} > r_{\text{Tabel}}$ (0.58) maka soal tersebut valid dan sebaliknya (Arikunto & Jabar, 2009). Teknik yang digunakan yaitu dengan korelasi product moment sebagai berikut:

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

Keterangan :

r_{xy} = koefisien validitas instrumen (korelasi antara X dan Y)

n = banyak responden

\bar{X} = Skor rata-rata X

\bar{Y} = Skor rata-rata Y

Reliabilitas instrumen diperoleh melalui item soal yang dikatakan valid berdasarkan hasil uji validitas empiris. Reliabilitas tes dapat dihitung dengan menggunakan persamaan cronbach alpha sebagai berikut:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right)$$

Keterangan :

r_{11} = reliabilitas yang dicari

$\sum \sigma_i^2$ = jumlah varians skor tiap-tiap item

σ_t^2 = varians total

Berdasarkan hasil analisis uji reliabilitas yang diperoleh, maka hasil analisis dikategorikan berdasarkan interval tingkat reliabilitas yang dapat dilihat pada Tabel 2.

Tabel 2. Kategori interval tingkat reliabilitas

Nilai peserta didik	Kriteria
$0,00 < r_{11} \leq 0,20$	Reliabilitas sangat rendah
$0,20 < r_{11} \leq 0,40$	Reliabilitas rendah
$0,40 < r_{11} \leq 0,60$	Reliabilitas sedang
$0,60 < r_{11} \leq 0,80$	Reliabilitas tinggi
$0,80 < r_{11} \leq 1,00$	Reliabilitas sangat tinggi

Sumber : (Hamzah, 2014)

III. Hasil dan Pembahasan

Pada pengembangan instrumen tes diagnostik *four-tier multiple choice* yang valid dan reliabel pada materi minyak bumi, maka pada penelitian melalui tahapan-tahapan menggunakan model pengembangan tipe *formative research* Tessmer (1993) yang meliputi tahap persiapan, penilaian sendiri, pembuatan prototipe (validasi, evaluasi dan revisi). Pembahasan hasil penelitian berdasarkan prosedur pengembangan tersebut dapat dijabarkan sebagai berikut:

Tahap persiapan (*preliminary*)

Pada tahap persiapan diawali dengan kajian beberapa referensi yang berkaitan dengan penelitian ini. Salah satu kajian referensi menunjukkan terdapat miskonsepsi pada konsep yang berkaitan dengan materi minyak bumi. Hasil kajian referensi menunjukkan adanya miskonsepsi pada materi minyak bumi meliputi konsep rumus struktur, reaksi pembakaran senyawa alkana dan reaksi perengkahan. Penyebab miskonsepsi yang terjadi adalah kemampuan dasar yang masih kurang serta minat belajar kimia yang rendah (Djarwo, 2013).

Selanjutnya dilakukan kegiatan penentuan tempat/lokasi penelitian. Lokasi pada penelitian ini adalah SMAN 2 Tanjungpinang sebagai lokasi peneliti melakukan praktik lapangan pengalaman yang belum menyediakan instrumen tes yang dikembangkan. Sedangkan subjek pada penelitian ini adalah 12 orang peserta didik kelas XI MIA. Setelah penentuan tempat dan subjek uji coba maka dilakukan wawancara yang bertujuan untuk mengidentifikasi kegiatan pembelajaran dan miskonsepsi peserta didik di SMA Negeri 2 Tanjungpinang. Wawancara dilakukan dengan guru kimia kelas XI MIA di sekolah tersebut. Berdasarkan hasil wawancara dengan guru kimia kelas XI SMAN 2 Tanjungpinang diketahui bahwa belum tersedianya instrumen tes yang didesain khusus untuk mengidentifikasi miskonsepsi dan kurangnya variasi instrumen tes yang digunakan. Peserta didik hanya diminta untuk mengemukakan alasan pilihan jawaban pada setiap soal pilihan ganda yang mereka kerjakan.

Tahap penilaian sendiri (*self evaluation*)

Tahapan ini bertujuan untuk merancang sebuah instrumen tes diagnostik *four-tier multiple choice* untuk mengidentifikasi miskonsepsi peserta didik pada materi minyak bumi untuk peserta didik SMA/MA. Tahapan ini terdiri dari dua tahap yaitu analisis dan desain. Pada tahap analisis, dilakukan analisis terhadap peserta didik dan juga materi.

Kegiatan analisis peserta didik difokuskan pada peserta didik kelas XI MIA SMA Negeri 2 Tanjungpinang sebagai subjek uji coba tahap satu-satu dan kelompok kecil. Pada uji satu-satu sebanyak 3 orang peserta didik yang terdiri dari peserta didik berkemampuan tinggi, sedang dan rendah. Sedangkan untuk kelompok kecil diambil 12 orang peserta didik yang terdiri dari 4 peserta didik berkemampuan tinggi, 4 berkemampuan sedang dan 4 peserta didik berkemampuan rendah. Subjek tersebut dipilih berdasarkan hasil belajar (Ulangan Harian) pada bab Hidrokarbon dan Minyak bumi. Nilai perolehannya dapat dilihat pada Tabel 3.

Tabel 3 Hasil ulangan bab hidrokarbon dan minyak bumi kelas XI MIA

Nama	Nilai	Kemampuan	Subjek
Peserta didik 1	91	Tinggi	Satu-satu
Peserta didik 2	78	Sedang	Satu-satu
Peserta didik 3	65	Rendah	Satu-satu
Peserta didik 4	93	Tinggi	Kelompok kecil
Peserta didik 5	90	Tinggi	Kelompok kecil
Peserta didik 6	88	Tinggi	Kelompok kecil
Peserta didik 7	87	Tinggi	Kelompok kecil
Peserta didik 8	78	Sedang	Kelompok kecil
Peserta didik 9	78	Sedang	Kelompok kecil
Peserta didik 10	75	Sedang	Kelompok kecil
Peserta didik 11	75	Sedang	Kelompok kecil
Peserta didik 12	65	Rendah	Kelompok kecil
Peserta didik 13	65	Rendah	Kelompok kecil
Peserta didik 14	50	Rendah	Kelompok kecil
Peserta didik 15	50	Rendah	Kelompok kecil

Selanjutnya dilakukan analisis terhadap materi. Adapun kompetensi dasar yang dianalisis yaitu KD 3.2 Menjelaskan proses pembentukan fraksi-fraksi minyak bumi, teknik pemisahan serta kegunaannya. Analisis keterkaitan tersebut dapat dilihat pada Tabel 4.

Tabel 4 Analisis keterkaitan KI, KD dan indikator

Kompetensi Inti	Indikator Pembelajaran	Materi/Submateri
KI 3: Memahami, menerapkan, dan menganalisis pengetahuan faktual, konseptual, prosedural, dan metakognitif berdasarkan rasa ingin tahunya tentang ilmu pengetahuan, teknologi, seni, budaya, dan humaniora dengan wawasan kemanusiaan, kebangsaan, kenegaraan, dan peradaban terkait penyebab fenomena dan kejadian, serta menerapkan pengetahuan prosedural	<ul style="list-style-type: none"> -Mengemukakan proses pembentukan minyak bumi menurut teori dupleks. -Menentukan eksplorasi minyak bumi -Menentukan teknik pemisahan -Mengurutkan teknik pemisahan pembentukan fraksi-fraksi minyak bumi -Menyimpulkan hasil tiap fraksi pada pengolahan tahap pertama minyak bumi -Menganalisis hasil tiap fraksi pada pengolahan tahap kedua minyak bumi beserta hasil akhirnya -Menentukan mutu bensin (angka oktan) -Menentukan cara meningkatkan mutu bensin -Menentukan proses pembakaran bensin pada mesin -Menentukan kegunaan senyawa hasil pengolahan minyak bumi 	<ul style="list-style-type: none"> -Proses terbentuknya minyak bumi -Eksplorasi minyak bumi - Teknik pemisahan -Destilasi sederhana -Teknik pemisahan destilasi bertingkat -Fraksi minyak bumi pengolahan tahap pertama -Fraksi minyak bumi pengolahan tahap kedua -Mutu bensin -Pembakaran produk minyak bumi

pada bidang kajian yang spesifik sesuai dengan bakat dan minatnya untuk memecahkan masalah	-Mengemukakan cara mengatasi dan dampak negatif penggunaan minyak bumi terhadap lingkungan dan kesehatan	-Kegunaan minyak bumi dalam kehidupan
--	--	---------------------------------------

Hasil analisis terhadap kompetensi dasar menghasilkan 11 indikator yang digunakan pada soal seperti yang disajikan pada Tabel 4. Indikator pencapaian kompetensi dikategorikan antara lain sebagai kompetensi minimal yang terdapat pada KD dan menggunakan kata kerja operasional yang bisa diukur (Ariyana dkk., 2018). Setelah kegiatan analisis materi dilakukan selanjutnya adalah merancang atau mendesain instrumen tes, meliputi: kisi-kisi tes, soal tes, kunci jawaban dan pedoman interpretasi dan penskoran. Tahapan awal yang dilakukan peneliti adalah merancang kisi-kisi tes berdasarkan indikator pencapaian dan ranah kognitif masing-masing soal sebagai bahan pertimbangan bagi validator untuk memeriksa validitas dari soal-soal tes diagnostik *four-tier*. Hasil pada tahap desain ini menghasilkan prototipe 1 yang akan dinilai oleh validator dan juga peserta didik tahap satu-satu.

Tahap pembuatan protipe (*prototyping*)

Tahap pembuatan prototipe ini bertujuan untuk menghasilkan prototipe II dari instrumen yang telah direvisi berdasarkan masukan dari para ahli dan data yang diperoleh dari uji coba satu-satu. Kegiatan pada tahap ini adalah penilaian ahli, satu-satu dan kelompok kecil. Kegiatan pada tahap ini meliputi validasi diikuti dengan revisi.

Validasi instrumen dilakukan dengan cara memberikan lembar validasi instrumen, kisi-kisi tes, soal tes dan pedoman interpretasi dan penskoran kepada validator. Dalam tahap validasi ini, validator menilai 13 indikator pada setiap item yang berkaitan dengan instrumen yang telah dirancang. Berdasarkan penilaian ketiga validator didapatkan penilaian untuk tiap item soal yang dapat dilihat pada Tabel 5.

Tabel 5 Penilaian validator pada setiap item

Item/no soal	V/Aspek			V/Item	Kategori
	Materi	Konstruksi	Bahasa		
1	0.88	0.85	0.88	0.87	Sangat valid
2	0.88	0.88	0.88	0.88	Sangat valid
3	0.88	0.85	0.88	0.87	Sangat valid
4	0.88	0.88	0.88	0.88	Sangat valid
5	0.94	0.87	0.88	0.90	Sangat valid
6	0.94	0.88	0.88	0.90	Sangat valid
7	0.94	0.88	0.88	0.90	Sangat valid
8	0.91	0.88	0.88	0.89	Sangat valid
9	0.88	0.87	0.88	0.88	Sangat valid
10	0.88	0.88	0.88	0.88	Sangat valid
11	0.88	0.88	0.88	0.88	Sangat valid
12	0.88	0.88	0.88	0.88	Sangat valid
13	0.94	0.88	0.88	0.90	Sangat valid
14	0.88	0.88	0.88	0.88	Sangat valid
15	0.88	0.88	0.88	0.88	Sangat valid
16	0.91	0.88	0.88	0.89	Sangat valid
17	0.88	0.87	0.88	0.88	Sangat valid
18	0.94	0.87	0.88	0.90	Sangat valid
19	0.94	0.88	0.88	0.90	Sangat valid
20	0.94	0.88	0.88	0.90	Sangat valid

Rata-rata	0.89	Sangat valid
-----------	------	--------------

Berdasarkan nilai hasil validasi ketiga validator, apabila dibandingkan dengan kriteria validitas yang diadaptasi dari Azwar (2012), maka skor pencapaian dari ketiga validator untuk keseluruhan item soal termasuk dalam kriteria sangat valid. Hasil validasi dari tiga validator tersebut menghasilkan validitas rata-rata sebesar 0.89 dengan kategori sangat valid. Kemudian instrumen tes *four-tier* direvisi berdasarkan komentar dan saran validator terhadap instrumen yang meliputi kisi-kisi tes, soal tes, dan kunci jawaban yang dapat dilihat pada Tabel 6.

Tabel 6 Komentar dan saran revisi validator

Validator	Saran revisi
Validator 1	<ol style="list-style-type: none"> 1. Pada soal no 1, sebaiknya penggunaan kalimat fotosintesa ganggang diganti dengan dekomposisi ganggang agar lebih sesuai dengan teori pembentukan minyak bumi 2. Soal no 3 batasi pertanyaan pada teori dupleks sehingga lebih mengarah ke jawaban yang diinginkan 3. Soal no 14 masih kurang sesuai dengan indikator soal yaitu cara meningkat mutu bensin
Validator 2	<ol style="list-style-type: none"> 1. Soal no 3 opsi A jawaban sebaiknya tidak perlu diperjelas cukup dengan menemukan daerah perangkap (trap) 2. Pada soal no 5 sebaiknya tabel berisi lebih dari 3 kategori fraksi. 3. Soal no 18 opsi tidak rasional, karena perbedaan kuantitas kata antara opsi satu dengan yang lain.
Validator 3	<ol style="list-style-type: none"> 1. Sesuaikan taksonomi bloom sehingga soal yang dibuat sesuai dengan tingkat kognitif yang diinginkan

Tahapan selanjutnya yang dilakukan pada proses pengembangan ini adalah tahap satu-satu, dengan melibatkan 3 orang peserta didik. Peserta didik yang dipilih merupakan peserta didik yang sebaya non subjek penelitian yang berkemampuan tinggi, sedang dan rendah berdasarkan nilai hasil belajar ketiga peserta didik tersebut. Ketiga peserta didik tersebut diminta berkomentar terhadap soal yang telah dikerjakan. Komentar yang diperoleh digunakan sebagai bahan pertimbangan merevisi desain instrumen tes yang telah dibuat, dari hasil revisi tinjauan ahli dan satu-satu ini menghasilkan prototipe II yang akan diuji cobakan pada kelompok kecil. Komentar peserta didik tahap satu-satu ini dapat dilihat pada Tabel 7.

Tabel 7 Komentar peserta didik tahap satu-satu

Peserta didik	Kemampuan	Komentar
Peserta didik 1	Tinggi	Perintah soal dan maksud soal sudah jelas serta diksi yang mudah dimengerti dan tidak bertele-tele
Peserta didik 2	Sedang	Soal yang diberikan sudah bagus, tetapi menurut saya tingkatan soal agak sulit sehingga saya sedikit kesulitan menjawabnya
Peserta didik 3	Rendah	Menurut saya soal yang diberikan sudah sangat jelas

Hasil dari penilaian ahli dan satu-satu dijadikan dasar untuk merevisi prototipe I menjadi desain prototipe II. Kemudian hasilnya diuji cobakan pada kelompok kecil sebanyak 12 orang peserta didik. Karakteristik peserta didik terdiri dari empat peserta didik dengan kemampuan tinggi, empat peserta didik dengan kemampuan sedang dan empat peserta didik dengan kemampuan rendah.

Berdasarkan hasil tes peserta didik inilah produk didapatkan validitas empiris dan reliabilitas. Adapun hasil analisis uji validitas empiris dapat dilihat pada Tabel 8.

Tabel 8 Hasil analisis validasi empiris

No. Soal	Uji validitas		
	r Hitung	r Tabel	Kriteria
1	0.68	0.58	Tinggi
2	0.65	0.58	Tinggi
3	0.85	0.58	Sangat tinggi
4	0.45	0.58	Cukup
5	0.71	0.58	Tinggi
6	0.73	0.58	Tinggi
7	0.69	0.58	Tinggi
8	0.61	0.58	Tinggi
9	0.68	0.58	Tinggi
10	0.64	0.58	Tinggi
11	0.75	0.58	Tinggi
12	0.79	0.58	Tinggi
13	0.64	0.58	Tinggi
14	0.65	0.58	Tinggi
15	0.62	0.58	Tinggi
16	0.61	0.58	Tinggi
17	0.62	0.58	Tinggi
18	0.35	0.58	Rendah
19	0.69	0.58	Tinggi
20	0.67	0.58	Tinggi

Hasil validitas empiris menghasilkan 1 soal dengan kriteria validitas sangat tinggi dengan r Hitung yaitu 0.85 dan 17 soal dengan kriteria validitas tinggi. Apabila dibandingkan dengan kriteria validitas empiris yang diadaptasi dari (Arikunto & Jabar, 2009), maka pencapaian ini termasuk dalam kriteria sangat tinggi dan tinggi. Pada saat penelitian terdapat 1 soal dengan kriteria cukup yaitu soal nomor 4, serta 1 soal dengan kriteria rendah yaitu soal nomor 18, kedua soal tersebut tergolong tidak valid dikarenakan banyak peserta didik yang tidak dapat menjawab soal tersebut sehingga menghasilkan nilai r Hitung sebesar 0.45 dengan kategori cukup atau tidak valid dan 0.35 dengan kategori rendah atau tidak valid. Apabila dibandingkan dengan kriteria validitas empiris, kategori tinggi berada pada rentang 0.60-0.80, kategori cukup 0.40-0.60, kategori rendah 0.20-0.40, hasil validitas empiris tiap butir diuji dengan korelasi *product moment* jika r Hitung > r Tabel (0.58) maka soal tersebut valid dan sebaliknya (Arikunto & Jabar, 2009). Maka dapat disimpulkan soal yang akan diuji reliabilitas sebanyak 18 Soal dengan kategori validitas sangat tinggi dan tinggi.

Uji reliabilitas berdasarkan hasil uji coba pada kelompok kecil yang melibatkan peserta didik kelas XI MIA sebanyak 12 orang. Berdasarkan hasil pekerjaan peserta didik maka dapat dihitung tingkat reliabilitas tes. Berdasarkan analisis data tersebut, reliabilitas tes yang diperoleh adalah 0.94 dengan interpretasi reliabilitas sangat tinggi dengan 18 Soal yang sudah valid. Adapun hasil analisis uji reliabilitas dapat dilihat pada Tabel 9.

Tabel 9 Hasil analisis uji reliabilitas

Σ Item	Jumlah varian butir	Varian total	r11
18	14.71	124.45	0.94

Reliabilitas instrumen tes pada uji coba kelompok kecil sebanyak 12 orang peserta didik secara umum dinyatakan reliabel, karena berdasarkan analisis uji reliabilitas diperoleh hasil sebesar 0.94 dengan kriteria reliabilitas sangat tinggi. Apabila dibandingkan dengan tabel kategori interval tingkat reliabilitas yang diadaptasi dari Hamzah (2014) yang menyatakan kategori tinggi pada rentang $0,60 < r_{11} \leq 0,80$ dan kategori sangat tinggi $0,80 < r_{11} \leq 1,00$, maka skor pencapaian ini termasuk dalam kriteria reliabilitas sangat tinggi.

IV. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa :

1. Instrumen tes diagnostik *four-tier multiple choice* untuk mengidentifikasi miskonsepsi peserta didik pada materi minyak bumi mempunyai langkah-langkah penelitian terdiri dari tahapan oleh model Tessmer (1993) yaitu: persiapan (*preliminary*), tahap penilaian sendiri (*self evaluation*), tahap pembuatan produk (*prototyping*) yang meliputi penilaian ahli, satu-satu, dan kelompok kecil.
2. Instrumen tes diagnostik *four-tier multiple choice* yang dikembangkan memiliki kriteria yang telah ditetapkan yaitu sangat valid pada validitas isi, penilaian oleh ketiga validator menghasilkan rata-rata validasi instrumen tes diagnostik sebesar 0.89 dengan kategori sangat valid. Sedangkan hasil dari validitas empiris yang diujikan pada tahap kelompok kecil menghasilkan jumlah soal dengan kategori validitas sangat tinggi sebanyak 1 soal, kemudian soal dengan kategori validitas tinggi sebanyak 17 dan 2 soal dengan kategori validitas cukup dan rendah atau tidak valid. Maka dapat disimpulkan sebanyak 18 Soal dengan kategori validitas sangat tinggi dan tinggi yang digunakan untuk uji reliabilitas.
3. Reliabilitas instrumen tes secara umum dinyatakan reliabel karena berdasarkan analisis instrumen tes reliabilitas yang diperoleh adalah 0.94 dengan interpretasi reliabilitas sangat tinggi.

V. Daftar Pustaka

- Ardiansah, A. (2016). Identifikasi Konsep Alternatif pada Guru Kimia: sebuah Kajian Literatur. Prosiding SNPS (Seminar Nasional Pendidikan Sains), 3, 49–54.
- Arikunto, S., & Jabar, C. S. A. (2009). Evaluasi Program Pendidikan. Jakarta: Bumi Aksara.
- Ariyana, Y., Pudjiastuti, A., Bestari, R., & Zamroni. (2018). Pembelajaran Berorientasi pada Keterampilan Berpikir Tingkat Tinggi. Kementerian pendidikan dan kebudayaan.
- Azwar, S. (2012). Reliabilitas dan Validitas Edisi 4. Yogyakarta: Pustaka Pelajar.
- Caleon, I. S., & Subramaniam, R. (2010). Do Students Know What They Know And What They Don't Know? Using A Four-Tier Diagnostic Test To Assess The Nature Of Students' Alternative Conceptions. *Research in Science Education*, 40(3), 313–337.
- Djarwo, C. F. (2013). Analisis Miskonsepsi Mahasiswa Pendidikan Kimia Pada Materi Hidrokarbon. *Jurnal Ilmiah IKIP Mataram*, 6(2), 90–97.
- Hamzah. (2014). Evaluasi Pembelajaran Matematika. Jakarta : Rajawali Press.
- Harahap, I. P. P., & Novita, D. (2020). Validitas Dan Reliabilitas Instrumen Tes Diagnostik Four-Tier Multiple Choice (4TMC) Pada Konsep Laju Reaksi. *Unesa Journal of Chemical Education*, 9(2), 222–227.
- Mahmudah, R. (2013). Identifikasi Miskonsepsi Peserta Didik Pada Konsep Suhu dan Kalor dengan Menggunakan Peta Konsep dan Wawancara. Yogyakarta: UIN Sunan Kalijaga.
- Suparno, P. (2013). Miskonsepsi & Perubahan Konsep Dalam Pendidikan Fisika. Jakarta: Gramedia Widiasarana.

Suwarto, D. (2013). Pengembangan Tes Diagnostik Dalam Pembelajaran. Yogyakarta: Pustaka Pelajar.

Tessmer, M. (1993). Planning and Conducting Formative Evaluations: Improving the Quality of Education and Training. Psychology Press.