

## PENGEMBANGAN TES DIAGNOSTIK *FOUR-TIER* UNTUK MENGIDENTIFIKASI MISKONSEPSI PESERTA DIDIK MATERI GELOMBANG DAN OPTIK

**Andry S. Utama Putra, Ida Hamidah, Nahadi**

Program Studi Magister Pendidikan IPA, Universitas Pendidikan Indonesia  
email: andrysutama@upi.edu

**Abstrak.** Penelitian bertujuan untuk mengembangkan tes diagnostik *four-tier* yang dapat digunakan untuk mengidentifikasi miskonsepsi peserta didik pada materi gelombang dan optik. Penelitian ini merupakan penelitian pengembangan dengan menggunakan model pengembangan Plomp yang dielaborasi dengan prosedur pengembangan tes diagnostik *four-tier* oleh Pesman dan Eryilmaz. Penelitian dilaksanakan di tiga SMP yang berlokasi di kecamatan Lembang pada peserta didik kelas VIII. Instrumen yang digunakan adalah lembar validasi ahli, lembar indentifikasi konsep alternatif peserta didik, dan instrumen validasi empiris. Analisis data yang digunakan yaitu analisis kuantitatif dengan melihat kualitas butir soal. Hasil penelitian menunjukkan bahwa uji validitas isi diperoleh nilai CVI sebesar 0,848 yang berarti termasuk dalam kriteria sangat sesuai. Uji validitas empiris diperoleh nilai outfit MNSQ 1,00 dan outfit ZSTD 0,0 yang berarti termasuk dalam kriteria nilai yang diterima. Sedangkan *Point Measure Correlation* keseluruhan butir soal sudah termasuk dalam rentang yang telah ditentukan. *Item reliability* 0.93 dan nilai *alpha Cronbach* 0,85 juga menunjukkan bahwa instrumen yang dikembangkan reliabilitasnya bagus. Nilai *person Measure*  $-0,29 < \logit$  0,0 menunjukkan bahwa tingkat kesulitan soal termasuk kategori sangat baik. Secara keseluruhan dapat disimpulkan bahwa tes diagnostik *four-tier* yang dikembangkan telah memenuhi kriteria dan dinyatakan layak digunakan untuk mengidentifikasi miskonsepsi peserta didik pada materi gelombang dan optik.

**Kata Kunci:** Tes diagnostik *four-tier*, miskonsepsi, gelombang dan optik

## DEVELOPMENT OF FOUR-TIER DIAGNOSTIC TESTS TO IDENTIFY MISCONCEPTION OF PARTICIPANTS IN WAVE AND OPTICAL MATERIALS

**Andry S. Utama Putra, Ida Hamidah, Nahadi**

Magister of Science Education Departement, Universitas Pendidikan Indonesia  
email: andrysutama@upi.edu

**Abstract.** The study aims to develop a four-tier diagnostic test that can be used to identify students' misconceptions in wave and optical material. This research is developing research using the Plomp development model that was elaborated by the procedure for developing a four-tier diagnostic test by Pesman and Eryilmaz. The study was conducted in three junior high schools located in the Lembang sub-district in class VIII students. The instruments used were expert validation sheets, identification sheets of alternative concepts for

students, and empirical validation instruments. Data analysis used is quantitative analysis by looking at the quality of the items. The results of the study showed that the content validity test obtained a CVI value of 0,848 which meant that the criteria were very suitable. The empirical validity test obtained the value of MNSQ outfit 1,00 and ZSTD outfit -0.0 which meant that it was included in the criteria of accepted values. While Point Measure Correlation, the whole item is included in the predetermined range. Item reliability is 0.93 and the alpha cronbach value of 0.85 also shows that the instrument developed has good reliability. The value of person Measure  $-0.29 < \text{logit } 0.0$  indicates that the level of difficulty of the question includes a very good category. Overall, it can be concluded that the four-tier diagnostic test developed has met the criteria and is declared feasible to be used to identify students' misconceptions in wave and optical material.

**Keywords:** *Four-tier diagnostic tests, misconceptions, waves and optics*

## PENDAHULUAN

Miskonsepsi dalam pembelajaran sains dikenal sebagai gagasan peserta didik yang tidak sesuai dengan gagasan yang diterima dan dipahami secara umum dalam pandangan ilmiah (Ilyas & Saeed, 2018; Faizah, 2016; Sanders & Makotsa, 2016). Miskonsepsi menjadi salah satu penyebab terjadinya kesulitan belajar yang berdampak pada rendahnya hasil belajar peserta didik (Mubarak, Susilaningsih, & Cahyono, 2016; Eryilmaz, 2010). Tidak adanya identifikasi sejak awal pada peserta didik akan menciptakan penghalang untuk restrukturisasi pengetahuan dan menyebabkan kegagalan dalam memahami suatu konsep (Zulfikar, Samsudin, & Saepuzaman, 2017; Vrabec & Proksa, 2016; Galvin, Simmie, & O'Grady, 2015).

Miskonsepsi pada pembelajaran sains yang terfokus pada bidang fisika sangat potensial terjadi karena beberapa konsep bersifat abstrak (Gurel, Eryilmaz, & McDermott, 2016). Disamping itu, penggunaan bahasa yang tidak baku menjadi salah satu kendala yang menyebabkan konsep fisika menjadi sulit dipahami oleh peserta didik (Kryjevskaja, Stetzer, & Heron, 2012).

Konsep fisika yang sering mengalami miskonsepsi adalah materi gelombang (Sutopo, 2016; I. Caleon & Subramaniam, 2010a) dan optik (Munawaroh & Setyarsih, 2016; Chu, Treagust, & Chandrasegaran, 2009). Materi gelombang dan optik memiliki hubungan yang tidak terpisahkan. Pemahaman terkait materi tersebut berdampak pada rekayasa teknologi yang mampu menciptakan gelombang-gelombang tertentu yang bisa dimanfaatkan untuk keperluan manusia, mulai dari gelombang radio, TV, *microwave*, *handphone* sampai gelombang sinar X dan sinar gamma (Gunawan, Harjono, & Sahidu, 2015). Pemahaman terkait materi gelombang dan optik juga sangat penting untuk dikuasai oleh peserta didik karena memiliki dampak dalam keberhasilan revolusi 4.0. Dijelaskan bahwa pemahaman terkait materi gelombang dan optik berdampak pada kemajuan teknologi fiber-optik, pita lebar, gawai/perangkat, dan pemanfaatan transmisi gelombang yang memiliki manfaat pada penyaluran internet keseluruh dunia (Pfeiffer, 2017). Berdasarkan penjelasan diatas mengindikasikan bahwa

pemahaman yang benar terkait materi gelombang dan optik sangatlah vital. Sehingga miskonsepsi yang terjadi pada materi tersebut harus dapat diminimalisir atau dihilangkan.

Miskonsepsi dapat muncul dari pengalaman sehari-hari yang dialami peserta didik ketika berinteraksi dengan lingkungan sekitarnya. Melalui pengalaman tersebut peserta didik akan membangun teori sendiri di dalam pikirannya yang belum tentu benar. Apabila intuisi yang terbentuk tidak benar, akan sangat sulit untuk diperbaiki karena tanpa sengaja secara konsisten konsep yang salah tersebut telah menjadi pegangan dan berdampak pada miskonsepsi yang berkelanjutan (Cil, 2015; Vitharana, 2015). Sehingga, dibutuhkan suatu metode untuk mengidentifikasi kemungkinan miskonsepsi yang dialami oleh peserta didik.

Ada beberapa metode yang digunakan untuk mengidentifikasi munculnya miskonsepsi dalam pembelajaran diantaranya pilihan ganda *two-tier* (Loh, Subramaniam, & Tan, 2014; Uyulgan, Akkuzu, & Alpat, 2014; Kanli, 2015; Korur, 2015; Pan & Chou, 2015; Yang, Chen, & Hwang, 2015), pilihan ganda *three-tier* (Milenkovic, Hrin, Segedinac, & Horvat, 2016; Taslidere, 2016), dan pilihan ganda *four-tier* (Gurel, Eryilmaz, & McDermott, 2017; Hermita et al., 2017; ; Zulfikar, Samsudin, & Saepuzaman, 2017; I. Caleon & Subramaniam, 2010a).

Metode pilihan ganda *two-tier*, pilihan ganda *three-tier*, dan pilihan ganda *four-tier* dikenal sebagai tes diagnostik. Tes diagnostik adalah tes yang digunakan untuk mengetahui kelemahan-kelemahan peserta didik sehingga berdasarkan kelemahan tersebut dapat diberikan perlakuan yang tepat (I. S. Caleon & Subramaniam, 2010b). Oleh karena itu, hasil tes diagnostik dapat digunakan sebagai landasan awal untuk memberikan tindak lanjut yang tepat dan sesuai dengan kelemahan yang dimiliki peserta didik.

Berdasarkan metode-metode yang telah disebutkan diatas, tes diagnostik *four-tier* merupakan metode yang paling mampu memberikan gambaran yang jelas mengenai miskonsepsi pada peserta didik (Gurel, Eryilmaz, & McDermott, 2015; Yang & Lin, 2015). Hasil dari

penelitian-penelitian tersebut memberikan informasi bahwa tes diagnostik *four-tier* dapat digunakan untuk mengidentifikasi miskonsepsi.

Khusus pada materi fisika, pengembangan tes diagnostik *four-tier* sudah cukup banyak dilakukan diantaranya pada materi listrik (Lestari, Suyana, & Jauhari, 2018), pada materi geometrik optik (Gurel, Eryilmaz, & McDermott, 2017), pada materi getaran (Zaleha, Samsudin, & Nugraha, 2017), serta pada materi impuls dan momentum (Diyanahesa, Kusairi, & Latifah, 2017). Akan tetapi kebanyakan tes diagnostik *four-tier* tersebut digunakan untuk mengidentifikasi miskonsepsi peserta didik pada jenjang SMA dan perguruan tinggi. Penggunaan tes diagnostik *four-tier* dalam mengidentifikasi miskonsepsi pada jenjang SMP belum banyak dilakukan. Miskonsepsi yang terjadi pada jenjang SMP akan sangat mempengaruhi pemahaman peserta didik dalam proses belajar pada jenjang berikutnya.

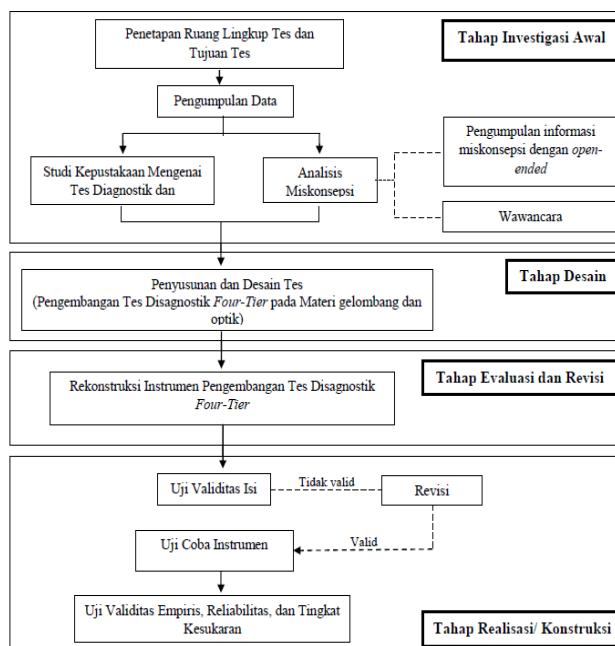
Adapun tujuan dari penelitian ini adalah mengembangkan tes diagnostik *four-tier* yang dapat digunakan untuk mengidentifikasi miskonsepsi peserta didik pada materi gelombang dan optik.

## METODE PENELITIAN

Penelitian yang digunakan adalah penelitian pengembangan (*Research and Development*) Plomp (Plomp & Nieveen, 2013) yang dielaborasi dengan prosedur pengembangan tes diagnostik *four-tier* oleh Pesman & Eryilmaz (2010). Pengembangan yang dilakukan berorientasi pada pengembangan produk. Penelitian dilaksanakan di tiga SMP yang berlokasi di kecamatan Lembang, kabupaten Bandung Barat, Jawa Barat pada peserta didik kelas VIII semester genap Tahun ajaran 2018/2019. Subjek penelitian dibagi menjadi dua yaitu untuk keperluan identifikasi konsep alternatif peserta didik dan subjek uji coba instrumen. Subjek penelitian untuk keperluan identifikasi konsep alternatif peserta didik dilakukan di SMPN 2 Lembang dengan total sampel berjumlah 45 peserta didik (kriterianya yaitu masing-masing 15 peserta didik dengan nilai ulangan harian tinggi, sedang, dan rendah). Informasi nilai ulangan harian diperoleh dari hasil studi dokumentasi nilai ulangan harian peserta didik dan wawancara dengan pendidik IPA yang bersangkutan. Untuk keperluan uji coba instrumen, digunakan sampel yang tersebar pada dua sekolah yaitu SMPN 3 Lembang dan SMPN 4 Lembang dengan total sampel 247 peserta didik.

Prosedur pengembangan terdiri dari lima tahapan yaitu 1) investigasi awal (*preliminary investigation*), 2) tahap desain (*the design phase*), 3) tahap realisasi/konstruksi (*the realization/construction*), 4) tahap evaluasi dan revisi (*the stage of evaluation and revision*), serta 5) tahap implementasi (*the stage of implementation*). Akan tetapi, pada penelitian ini dibatasi hanya sampai pada tahap evaluasi dan revisi. Tahap implementasi tidak dilaksanakan karena tujuan utama penelitian adalah mengembangkan produk berupa tes diagnostik *four-tier*.

Prosedur penelitian disajikan dalam bentuk bagan di bawah ini.



Gambar 1. Prosedur Penelitian Pengembangan Plomp Yang Dielaborasi dengan Prosedur Pengembangan Tes Diagnostik *Four-Tier*

Instrumen penelitian yang digunakan dalam penelitian ini yaitu 1) instrumen lembar validasi ahli. Instrumen ini bertujuan mengetahui kriteria kevalidan tes diagnostik yang dikembangkan. Lembar validasi ini akan diberikan kepada ahli dan reviewer yang berkompeten dalam menilai dan memberi saran; 2) Instrumen indentifikasi konsep alternatif peserta didik (*two-tier open ended*). Instrumen ini digunakan untuk mengumpulkan jawaban alternatif peserta didik mengenai materi gelombang dan optik. Jawaban tersebut akan ditabulasi kemudian di analisis alasan-alasan yang teridentifikasi miskonsepsi. Jawaban peserta didik yang teridentifikasi miskonsepsi akan dijadikan sebagai pilihan alasan pada *tier* ketiga dalam perangkat tes diagnostik *four-tier*; dan 3) instrumen validasi empiris. Instrumen ini merupakan transformasi dari instrumen *two-tier open ended* menjadi tes diagnostik *four-tier* dan hasil revisi dari ahli dan *reviewer*.

Analisis data yang digunakan yaitu analisis kuantitatif dengan melihat kualitas butir soal yang meliputi uji validitas, reliabilitas, dan tingkat kesulitan soal. Uji validitas yang dilakukan terbagi menjadi 2 yaitu uji validitas berdasarkan isi dan uji validitas berdasarkan struktur internal (validitas empiris). Uji validitas berdasarkan isi diperoleh melalui *judgement* ahli dan *reviewer* kemudian melihat *Content Validity Ratio* (CVR) dan *Content Validity Index* (CVI). Perhitungan nilai CVR dilakukan berdasarkan persamaan Lawshe (1975) sebagai berikut :

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

Keterangan:

$n_e$  = Jumlah validator yang menyatakan valid  
 $N$  = Jumlah total validator

Kriteria nilai CVR minimum untuk jumlah *judgement* enam orang adalah lebih besar atau sama dengan 0,672 ( $CVR \geq 0,672$ ). Selain itu, dilakukan pula penentuan validitas tes secara keseluruhan dengan menggunakan metode *Content Validity Index* (CVI).

Tabel 1. Kriteria Hasil Perhitungan CVI

No.	Rentang	Kategori
1.	$CVI \geq 0,68$	Sangat Sesuai
2.	$0,34 \leq CVI < 0,67$	Sesuai
3.	$CVI < 0,34$	Tidak Sesuai

(Wilson, Pan, & Schumsky, 2012)

Uji validitas berdasarkan struktur internal (validitas empiris), reliabilitas, dan dan tingkat kesulitan soal dianalisis menggunakan model Rasch. Keunggulan dari model Rasch dibanding metode lainnya yaitu kemampuan melakukan prediksi terhadap data yang hilang (*missing data*), yang didasarkan pada pola respon yang sistematis. Dengan demikian data yang diperoleh seolah-olah sebagai data yang lengkap dan lebih akurat dalam analisis statistik selanjutnya (Sumintono & Widhiarso, 2015).

Uji validitas berdasarkan struktur internal (validitas empiris) dilakukan dengan bantuan program *winsteps*. Hal yang dilihat adalah berdasarkan kriteria nilai *Outfit Mean Square (MNSQ)* yaitu  $0,5 < MNSQ < 1,5$ ; nilai *Outfit Z-Standard (ZSTD)* yaitu  $-2,0 < ZSTD < +2,0$ ; dan nilai *Point Measure Correlation (Pt Mean Corr)* yaitu  $0,2 < Pt Mean Corr < 0,85$  (Boone, Staver, & Yale, 2014). Tingkat kesulitan soal didasarkan pada perbandingan *person measure* dan *item measure*. Jika *person measure* > *item measure* maka tingkat kesulitan soal dalam kategori baik (Sumintono & Widhiarso, 2015). Sedangkan untuk reliabilitas instrumen dianalisis dengan cara membandingkan nilai *person reliability*, *item reliability*, dan nilai *alpha Cronbach* yang diperoleh dengan kriteria koefisien reliabilitas, sebagai berikut:

Tabel 2. Kriteria Reliabilitas

No.	Rentang	Kategori
1.	$0,94 \leq r < 1,00$	Reliabilitas istimewa
2.	$0,91 \leq r < 0,94$	Reliabilitas bagus sekali
3.	$0,80 \leq r < 0,91$	Reliabilitas bagus
4.	$0,67 \leq r < 0,80$	Reliabilitas cukup
5.	$0,00 \leq r < 0,67$	Reliabilitas lemah

(Fisher, 2007)

## HASIL DAN PEMBAHASAN

Pengembangan tes diagnostik *four-tier* yang dilakukan menggunakan model pengembangan Plomp (Plomp & Nieveen, 2013) yang dielaborasi dengan prosedur pengembangan tes diagnostik *four-tier* oleh Pesman & Eryilmaz (2010). Penelitian ini hanya dibatasi sampai tahap pengembangan atau hanya sampai tahap ke empat dari prosedur pengembangan tes diagnostik secara keseluruhan. Berikut penjelasan terkait prosedur pengembangan tes diagnostik yang dilakukan:

Tahap investigasi awal (*preliminary investigation*). Tahap ini merupakan tahapan dasar yang dilakukan dalam mengembangkan tes diagnostik *four-tier*. Pada tahap ini dilakukan penetapan ruang lingkup tes dan tujuan tes melalui analisis silabus (KI dan KD) SMP yang terdapat pada kurikulum 2013. Berdasarkan hasil analisis diperoleh beberapa label konsep yang akan dikembangkan menjadi butir soal untuk mengidentifikasi miskonsepsi pada materi gelombang dan optik. Analisis terhadap kurikulum 2013 bertujuan untuk menentukan indikator yang mewakili setiap label konsep.

Tabel 3. Label konsep dan indikator pembelajaran sebagai acuan penyusunan tes diagnostik *four-tier*

Label Konsep	Indikator
Getaran	Menganalisis konsep getaran pada peristiwa ayunan bandul
	Menentukan besaran fisika pada gelombang transversal
	Menjelaskan pengertian gelombang transversal
Gelombang	Menentukan besaran fisika pada gelombang longitudinal
	Menganalisis sifat gelombang mekanik
	Menganalisis hubungan antara panjang gelombang, frekuensi, cepat rambat, dan periode gelombang
	Menjelaskan medium rambatan bunyi
	Menganalisis medium rambatan bunyi
Bunyi	Menjelaskan pengaruh amplitudo terhadap bunyi
	Menganalisis karakteristik bunyi
	Menjelaskan proses resonansi
	Menjelaskan penerapan frekuensi ultrasonik
Sifat-sifat cahaya	Menganalisis proses pemantulan cahaya
	Menganalisis proses pembiasan cahaya
Pembentukan bayangan	Menentukan pembentukan bayangan pada lensa
	Menentukan pembentukan bayangan pada cermin
	Menganalisis proses terbentuknya bayangan pada penderita cacat mata
Alat optik	Menentukan pembentukan bayangan pada alat optik

Selanjutnya proses pengumpulan data dilakukan melalui studi kepustakaan dan analisis miskonsepsi. Hasil dari studi kepustakaan mengenai tes diagnostik *four-tier* dan miskonsepsi pada materi gelombang dan optik memberikan gambaran terkait miskonsepsi-miskonsepsi yang sering dialami oleh peserta didik pada materi tersebut. Berikut beberapa miskonsepsi yang sering dialami peserta didik disajikan pada Tabel 4.

Tabel 4. Miskonsepsi Peserta Didik Pada Materi Gelombang Dan Optik

Materi	Miskonsepsi Peserta Didik
Gelombang	Semakin besar amplitudo, maka nada bunyi akan semakin lemah.
	Besar periode berbanding lurus dengan besar frekuensi.
	Bunyi merambat paling cepat pada medium udara.
	Cepat rambat bunyi selalu konstan untuk setiap keadaan.
	Bunyi bergerak sebagai gelombang transversal, karena tegak lurus dengan arah rambatannya.
	Semakin keras bunyi maka laju bunyi akan semakin cepat.
Optik	Bunyi tidak dapat merambat melalui zat padat dan zat cair.
	Lampu senter menggunakan cermin cembung karena cermin cembung bersifat dapat menyebarkan berkas cahaya.
	Lensa membentuk bayangan dengan cara memantulkan sinar.
	Sifat bayangan yang dibentuk oleh cermin datar adalah nyata dan terbalik.
	Ukuran bayangan tidak bergantung pada sudut pada mata yang berhadapan dengan objeknya
	Semakin redup cahaya, maka bayangan yang dibentuk akan semakin besar.
	Bayangan maya dan nyata merupakan bayangan yang dapat dilihat secara langsung oleh mata

Analisis miskonsepsi pada tahap investigasi awal (*preliminary investigation*) dilakukan untuk memperoleh informasi mengenai konsep alternatif peserta didik. Subjek penelitian pada tahap ini dilakukan di SMPN 2 Lembang dengan total sampel berjumlah 45 peserta didik (kriterianya yaitu masing-masing 15 peserta didik dengan nilai ulangan harian tinggi, sedang, dan rendah). Pada proses tersebut peserta didik diberikan soal dalam bentuk pilihan ganda beralasan (tes diagnostik *two-tier open ended*) sebanyak 33 soal. Berdasarkan soal tersebut peserta didik diminta memilih jawaban yang benar dan menuliskan alasan memilih jawaban tersebut. Jawaban yang diberikan peserta didik

akan ditabulasi untuk selanjutnya digunakan untuk pilihan alasan pada tes diagnostik *four-tier* yang dikembangkan. Selain itu juga dilakukan proses wawancara terhadap pendidik IPA terkait miskonsepsi yang sering dialami oleh peserta didik terkait materi gelombang dan optik. Informasi melalui proses wawancara dijadikan sebagai informasi tambahan dalam penyusunan instrumen tes diagnostik *four-tier* yang dikembangkan.

<b>Soal ke-n</b> (Butir Soal Pilihan Ganda).
<b>n.1 (Pilihan Jawaban).</b> A. (pilihan jawaban 1) B. (pilihan jawaban 2) C. (pilihan jawaban 3) D. (pilihan jawaban 4)
<b>n.2 Alasan terkait pilihan jawaban:</b> ..... .....

Gambar 2. Desain Tes Diagnostik *Two-Tier Open Ended*

Tahap kedua yaitu tahap desain (*the design phase*). Pada tahap ini hasil dari identifikasi miskonsepsi pada tahap investigasi awal (*preliminary investigation*) akan dianalisis untuk selanjutnya dikembangkan menjadi butir soal. Identifikasi miskonsepsi akan melahirkan sebuah pola yang terdiri dari konsep yang benar dan konsep yang terdapat miskonsepsi di dalamnya. Berdasarkan pola tersebut akan dikembangkan menjadi sebuah tes diagnostik berbentuk pilihan ganda tipe *four-tier*. Pada tingkat pertama terdiri empat pilihan jawaban, tingkat ketiga terdiri dari empat pilihan alasan, tingkat kedua dan keempat merupakan tingkat keyakinan terhadap jawaban yang diberikan pada tingkat pertama dan ketiga. Tingkat keyakinan ini terdiri dari dua pilihan yaitu yakin dan tidak yakin (Gurel, Eryilmaz, & McDermott, 2017).

<b>Soal ke-n</b> (Butir Soal Pilihan Ganda).
<b>n.1 (Pilihan Jawaban).</b> A. (pilihan jawaban 1) B. (pilihan jawaban 2) C. (pilihan jawaban 3) D. (pilihan jawaban 4)
<b>n.2 Tingkat keyakinan terkait pilihan jawaban</b> A. Yakin B. Tidak Yakin
<b>n.3 Pilihan alasan terkait jawaban yang dipilih</b> A. (pilihan alasan 1) B. (pilihan alasan 2) C. (pilihan alasan 3) D. (pilihan alasan 4)
<b>n.4 Tingkat keyakinan pilihan alasan</b> A. Yakin B. Tidak Yakin

Gambar 3. Desain Tes Diagnostik *Four-Tier* Yang Dikembangkan

Tahap ketiga dalam penelitian ini yaitu tahap realisasi/konstruksi (*the realization/construction*). Setelah membuat desain instrumen tes diagnostik *four-tier*, peneliti mengkonstruksi desain instrumen tes diagnostik tersebut. Konstruksi instrumen tes diagnostik *four-tier* yang dikembangkan disesuaikan dengan desain yang telah dijelaskan pada tahap sebelumnya. Konstruksi tersebut merupakan bentuk realisasi instrumen tes diagnostik *four-tier* yang digunakan oleh peneliti untuk validasi empiris. Instrumen tes diagnostik *four-tier* terdiri dari 33 soal. Salah satu contoh konstruksi tes diagnostik *four-tier* yang dikembangkan, disajikan pada Gambar 4.

<b>Pertanyaan</b>
1. Sebuah benda diletakkan di depan cermin cekung dengan fokus 2 cm. Jika benda berada pada jarak 4 cm di depan cermin, maka sifat bayangan yang terbentuk adalah....
<b>Pilihan Jawaban (Tingkat I)</b>
A. nyata, terbalik, dan diperbesar B. nyata, terbalik, dan diperkecil C. maya, tegak, dan diperbesar D. nyata, terbalik, dan sama besar
<b>Tingkat Keyakinan Terhadap Pilihan Jawaban (Tingkat II)</b>
A. Ya B. Tidak
<b>Alasan Berdasarkan Pilihan Jawaban (Tingkat III)</b>
A. benda terletak tepat di titik fokus cermin dan bayangan terletak di belakang cermin B. benda terletak di ruang II dan bayangan terletak di depan cermin C. benda terletak di ruang III dan bayangan terletak di depan cermin D. benda terletak tepat di titik pusat kelengkungan cermin dan bayangan terletak di depan cermin
<b>Tingkat Keyakinan Terhadap Pilihan Alasan (Tingkat IV)</b>
A. Ya B. Tidak

Gambar 4. Salah Satu Contoh Kontruksi Soal Tes Diagnostik *Four-Tier*

Selanjutnya tahap evaluasi dan revisi (*the stage of evaluation and revision*). Tahap ini dilakukan dengan melakukan uji kualitas tes diagnostik *four-tier* yang dikembangkan. Tahap evaluasi dan revisi dilakukan melalui 3 tahapan yaitu uji validitas (uji validitas berdasarkan isi dan uji validitas berdasarkan struktur internal), reliabilitas, dan tingkat kesulitan soal. Uji validitas berdasarkan isi dilakukan dengan melihat hasil *judgement* ahli dan *reviewer*. *Judgement* yang digunakan pada penelitian ini berjumlah 3 orang yang merupakan ahli dalam bidang asesmen dan bidang fisika. Sedangkan *reviewer* pada penelitian ini merupakan pendidik IPA pada setiap sekolah yang digunakan sebagai tempat penelitian. Hasil dari *judgement* tersebut kemudian dianalisis

menggunakan metode CVR (*Content Validity Ratio*) dan *Content Validity Index* (CVI). Pendekatan validitas berdasarkan isi yang dilakukan untuk mengetahui kesesuaian item dengan domain yang diukur (didasarkan pada aspek materi, konstruksi soal, dan bahasa).

Berdasarkan analisis yang dilakukan diperoleh nilai CVR dan CVI sebesar 0,848 yang berarti telah memenuhi kriteria minimum CVR dan termasuk kriteria sangat sesuai. Hal tersebut menunjukkan bahwa instrumen tes diagnostik *four-tier* yang dikembangkan dinyatakan valid dan dapat digunakan dalam tahap uji coba untuk melihat validitas berdasarkan struktur internal (validitas empiris), reliabilitas, dan tingkat kesulitan soal. Dalam proses uji validitas berdasarkan isi terdapat beberapa soal yang direvisi dan dihilangkan karena beberapa alasan, sehingga soal yang dinyatakan valid dan dapat digunakan dalam tahap uji coba sebanyak 28 soal.

Tabel 5. Hasil *judgement* terkait soal yang direvisi dan dihilangkan

Nomor Soal	Saran
2	Redaksi soal diperbaiki. Kalimat “masing-masing ujung statif diikat tali” dihilangkan karena dianggap kalimat yang ambigu.
7	Redaksi soal diperbaiki. Kalimat “perhatikan gambar berikut” dilengkapi dan diperjelas sehingga menjadi “perhatikan gambar pembentukan gelombang pada pegas”.
11	Redaksi soal diperbaiki. Kalimat “jarak yang sama dari mereka berdua” diubah menjadi “Tono dan Toni memiliki jarak yang sama terhadap Bayu”.
13	Option jawaban disusun secara hierarki.
17	Redaksi soal diperbaiki. Kalimat “Manakah pernyataan yang benar” diubah menjadi “Pernyataan yang paling sesuai”.
18	Redaksi kata pada option jawaban dihilangkan karena dianggap ambigu. Kata tersebut adalah “resonansi”.
19	Redaksi soal diperbaiki. Kalimat “Menurut kamu mengapa manusia tidak dapat mendengar” diubah menjadi “Penyebab manusia tidak dapat mendengar bunyi”.
25	Option jawaban A dan D diubah karena tidak berfungsi sebagai distractor.
28	Option jawaban B dan D diubah karena tidak berfungsi sebagai distractor.
24	Redaksi kata pada option jawaban dihilangkan. Kata tersebut adalah “Proses”.
5, 20, 21, 29, dan 31	Soal dihilangkan karena tidak mencirikan tes diagnostik/tingkat kesulitan soal.

Instrumen tes diagnostik *four-tier* yang telah direvisi dan dinyatakan valid oleh *judgement* kemudian di uji coba pada peserta didik kelas VIII di SMPN 3

Lembang (terdiri dari 120 peserta didik) dan SMPN 4 Lembang (terdiri dari 127 peserta didik). Peserta didik dalam tahap uji coba ini telah mempelajari materi terkait gelombang dan optik. Analisis data yang dilakukan pada tahap ini menggunakan model Rasch program *winsteps*. Berikut hasil analisis data menggunakan model Rasch disajikan pada Tabel 6 dan Tabel 7.

Tabel 6. Ringkasan statistik instrumen terkait validitas, reliabilitas, dan tingkat kesulitan soal

SUMMARY OF 247 MEASURED PERSON									
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT		
					MISQ	ZSTD	MISQ	ZSTD	
MEAN	15.7	28.0	-.29	.41	1.00	.0	1.00	.0	
S.D.	4.2	.0	.70	.03	.11	.9	.16	.9	
MAX.	24.0	28.0	1.89	.55	1.38	2.9	1.69	3.0	
MIN.	4.0	28.0	-1.89	.39	.78	-2.3	.70	-2.2	
REAL RMSE	.42	TRUE SD	.55	SEPARATION	1.30	PERSON RELIABILITY	.63		
MODEL RMSE	.42	TRUE SD	.56	SEPARATION	1.35	PERSON RELIABILITY	.65		
S.E. OF PERSON MEAN = .04									
PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00									
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .85									
SUMMARY OF 28 MEASURED ITEM									
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT		
					MISQ	ZSTD	MISQ	ZSTD	
MEAN	138.7	247.0	.00	.14	1.00	.0	1.00	.0	
S.D.	28.0	.0	.54	.01	.07	1.6	.09	1.6	
MAX.	200.0	247.0	-1.73	.17	.96	-.4	.88	-.9	
MIN.	96.0	247.0	1.68	.14	1.01	.3	1.03	.7	
REAL RMSE	.14	TRUE SD	.52	SEPARATION	3.69	ITEM RELIABILITY	.93		
MODEL RMSE	.14	TRUE SD	.52	SEPARATION	3.74	ITEM RELIABILITY	.93		
S.E. OF ITEM MEAN = .10									

Tabel 7. Statistik instrumen terkait *Point Measure Correlation*

	PT-MEASURE CORR.	EXACT EXP.	MATCH OBS%	ITEM EXP%	ITEM
A	.33	.31	51.0	63.2	127
B	.28	.28	78.1	78.0	13
C	.37	.31	62.3	65.1	16
D	.24	.30	67.2	71.4	15
E	.32	.31	58.7	62.9	112
F	.42	.31	62.3	64.1	14
G	.35	.31	61.9	62.9	12
H	.26	.31	63.6	67.6	11
I	.47	.32	60.3	64.0	17
J	.48	.30	66.4	65.2	128
K	.40	.31	64.0	63.4	120
L	.27	.31	64.0	63.6	121
M	.30	.31	67.6	68.5	113
N	.32	.31	66.0	65.0	118
n	.31	.31	69.2	68.1	125
m	.34	.31	67.2	64.8	114
l	.34	.31	65.6	63.2	115
k	.39	.31	72.9	68.8	126
j	.35	.30	72.9	72.7	19
i	.37	.31	65.6	63.0	123
h	.35	.27	80.2	80.9	116
g	.38	.31	66.4	63.4	124
f	.39	.31	66.4	63.6	122
e	.42	.31	66.8	63.0	18
d	.43	.31	72.1	69.0	110
c	.44	.30	71.7	70.5	119
b	.46	.31	71.3	62.8	111
a	.47	.31	70.4	63.8	117
			66.9	66.5	
			5.7	4.6	

Data yang digunakan untuk melihat validitas berdasarkan struktur internal (validitas empiris) secara berturut-turut yaitu outfit MNSQ, outfit ZSTD, dan *Point Measure Correlation (Pt Mean Corr)*. Berdasarkan Tabel 6 dapat ditunjukkan bahwa nilai outfit MNSQ yang diperoleh yaitu 1,00, nilai tersebut termasuk dalam kriteria nilai outfit MNSQ yang diterima (makin mendekati 1,00 maka kualitas butir soal semakin baik). Begitupula nilai outfit ZSTD yang diperoleh yaitu 0,0, nilai tersebut juga termasuk dalam kriteria nilai outfit ZSTD yang diterima (makin mendekati nilai 0,0 maka kualitas butir soal semakin baik). Sedangkan untuk *Point Measure Correlation (Pt Mean Corr)*, berdasarkan Tabel 7 menunjukkan bahwa keseluruhan butir soal sudah termasuk dalam rentang yang telah ditentukan. Hal tersebut berarti seluruh butir soal pada instrumen yang dikembangkan dapat mengukur apa yang seharusnya diukur dan tidak ada butir soal yang harus dihilangkan (Bond & Fox, 2015; Boone, Staver, & Yale, 2014). Secara keseluruhan instrumen tes diagnostik *four-tier* yang dikembangkan dinyatakan valid berdasarkan uji validitas empiris

Berdasarkan Tabel 6, diperoleh nilai *person reliability* 0.63 dan *item reliability* 0.93 yang menunjukkan bahwa konsistensi jawaban dari peserta didik lemah, namun kualitas butir-butir soal dalam tes diagnostik *four-tier* yang dikembangkan reliabilitasnya bagus. Sedangkan nilai *alpha Cronbach* 0,85 juga menunjukkan bahwa instrumen yang dikembangkan reliabilitasnya bagus dalam hal interaksi antara *person* dan butir-butir soal secara keseluruhan. Kedua analisis reliabilitas tersebut menunjukkan bahwa secara keseluruhan tes diagnostik *four-tier* yang dikembangkan dianggap mempunyai level konsisten yang tinggi sehingga dapat digunakan dalam penelitian yang sebenarnya karena telah memenuhi persyaratan (Sumintono & Widhiarso, 2015).

*Person Measure* -0,29 menunjukkan rata-rata nilai seluruh peserta didik dalam mengerjakan butir-butir soal yang diberikan. Nilai rata-rata yang lebih kecil dari nilai *logit* 0,0 menunjukkan kecenderungan abilitas peserta didik yang lebih kecil daripada tingkat kesulitan soal. Artinya tingkat kesulitan soal dalam instrumen *four-tier* yang dikembangkan sangat baik.

Setelah melalui keseluruhan tahapan pengembangan, tes diagnostik *four-tier* yang dikembangkan dikategorikan layak digunakan untuk mengidentifikasi miskonsepsi peserta didik pada materi gelombang dan optik.

## KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan dapat disimpulkan bahwa tes diagnostik *four-tier* yang dikembangkan telah memenuhi kriteria uji validitas (validitas berdasarkan isi dan validitas berdasarkan struktur internal), uji reliabilitas, dan uji tingkat kesulitan soal. Hal tersebut berarti tes diagnostik *four-tier* yang

dikembangkan dinyatakan layak digunakan untuk mengidentifikasi miskonsepsi peserta didik pada materi gelombang dan optik.

Saran untuk penelitian selanjutnya yaitu melakukan implementasi tes diagnostik *four-tier* yang telah dikembangkan, kemudian untuk proses analisisnya dapat menggunakan keputusan pola jawaban peserta didik pada setiap tingkatan. Hasil implementasi tes diagnostik *four-tier* tersebut dapat digunakan oleh pendidik sebagai dasar untuk menentukan tindak lanjut yang tepat dalam meminimalisir atau bahkan menghilangkan miskonsepsi yang terjadi pada peserta didik terkait materi gelombang dan optik.

## DAFTAR PUSTAKA

- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model Fundamental Measurement in the Human Sciences (3rd edition)*. New York: Routledge.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. New York: Springer Dordrecht.
- 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.
- 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, 32*(7), 939–961.
- Chu, H., Treagust, D. F., & Chandrasegaran, A. L. (2009). A Stratified Study of Students' Understanding of Basic Optics Concepts in Different Contexts Using Two-Tier Multiple-Choice Items. *Research in Science & Technological Education, 27*(3), 253–265.
- Cil, E. (2015). Effect of Two-tier Diagnostic Tests on Promoting Learners' Conceptual Understanding of Variables in Conducting Scientific Experiments. *Applied Measurement in Education, 28*(4), 253–273.
- Diyanahesa, N. E.-H., Kusairi, S., & Latifah, E. (2017). Development of Misconception Diagnostic Test in Momentum and Impulse Using Isomorphic Problem. *Journal of Physics: Theories and Applications, 1*(2), 145.
- Eryilmaz, A. (2010). Development and Application of Three-Tier Heat and Temperature Test: Sample of Bachelor and Students Graduate. *Eurasian Journal of Educational Research, 40*(40), 53–76.
- Faizah, K. (2016). Miskonsepsi Dalam Pembelajaran IPA. *Jurnal Darussalam: Jurnal Pendidikan Komunikasi dan Pemikiran Hukum Islam, VIII*(1), 115–128.
- Fisher, W. P. J. (2007). Rating Scale Instrument Quality Criteria. Diambil dari <http://www.rasch.org/rmt/rmt211a.htm>.
- Galvin, E., Simmie, G. M., & O'Grady, A. (2015). Identification of Misconceptions in the Teaching of Biology: A Pedagogical Cycle of Recognition, Reduction and Removal. *Higher Education of Social Science, 8*(2), 1–8.
- Gunawan, Harjono, A., & Sahidu, H. (2015). Pengembangan Model Laboratorium Virtual Berorientasi Pada Kemampuan Pemecahan Masalah Bagi Calon Guru Fisika. *Jurnal Materi dan Pembelajaran Fisika (JMPPF), 5*(2), 41–46.
- 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.
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2016). Identifying Pre-Service Physics Teachers' Misconceptions and Conceptual Difficulties about Geometrical Optics. *European Journal of Physics, 37*(4), 1–30.
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2017). Development and Application of a Four-Tier Test to Assess Pre-Service Physics Teachers' Misconceptions about Geometrical Optics. *Research in Science and Technological Education, 35*(2), 238–260.
- Hermita, N., Suhandi, A., Syaodih, E., Samsudin, A., Isjoni, Johan, H., Safitri, D. (2017). Constructing and Implementing a Four Tier Test about Static Electricity to Diagnose Pre-service Elementary School Teacher' Misconceptions. *Journal of Physics: Conference Series, 895*(1).
- Ilyas, A., & Saeed, M. (2018). Exploring Teachers' Understanding about Misconceptions of Secondary Grade Chemistry Students. *International Journal for Cross Disciplinary Subjects in Education (IJCDSE), 9*(1), 3323–3328.
- Kanli, U. (2015). Using a Two-Tier Test to Analyse Students' and Teachers' Alternative Concepts in Astronomy. *Science Education International, 26*(2), 148–165.
- Korur, F. (2015). Exploring Seventh-Grade Students' and Pre-service Science Teachers' Misconceptions in Astronomical Concepts. *Eurasia Journal of Mathematics, Science and Technology Education, 11*(5), 1041–1060.
- Kryjevskaja, M., Stetzer, M. R., & Heron, P. R. L. (2012). Student Understanding of Wave Behavior at a Boundary: The Relationships Among Wavelength, Propagation Speed, and Frequency. *American Journal of Physics, 80*(4), 339–347.
- Lawshe, C. H. (1975). A Quantitative Approach to Content Validity. *Personnel Psychology, 28*(4), 563–727.
- Lestari, I. N. M., Suyana, I., & Jauhari, A. (2018). Pengembangan Electricity Concept Test Berformat Four-Tier Test. *Jurnal Wahana Pendidikan Fisika, 3*(1), 69–73.
- Loh, A. S. L., Subramaniam, R., & Tan, K. C. D. (2014). Exploring Students' Understanding of



- Electrochemical Cells Using an Enhanced Two-Tier Diagnostic Instrument. *Research in Science and Technological Education*, 32(3), 229–250.
- Milenkovic, D. D., Hrin, T. N., Segedinac, M. D., & Horvat, S. (2016). Development of a Three-Tier Test as a Valid Diagnostic Tool for Identification of Misconceptions Related to Carbohydrates. *Journal of Chemical Education*, 93(9), 1514–1520.
- Mubarak, S., Susilaningih, E., & Cahyono, E. (2016). Pengembangan Tes Diagnostik Three Tier Multiple Choice untuk Mengidentifikasi Miskonsepsi Peserta Didik Kelas XI. *Journal of Innovative Science Education*, 5(2), 101–110.
- Munawaroh, R., & Setyarsih, W. (2016). Identifikasi Miskonsepsi Siswa dan Penyebabnya pada Materi Alat Optik Menggunakan Three-tier Multiple Choice Diagnostic Test. *Jurnal Inovasi Pendidikan Fisika*, 05(02), 79–81.
- Pan, S. J. A., & Chou, C. (2015). Using a Two-Tier Test to Examine Taiwanese Graduate Students' Misunderstanding of Responsible Conduct of Research. *Ethics and Behavior*, 25(6), 500–527.
- Pesman, H., & Eryilmaz, A. (2010). Development of a Three-Tier Test to Assess Misconceptions about Simple Electric Circuits. *Journal of Educational Research*, 103(3), 208–222.
- Pfeiffer, S. (2017). The Vision of “Industrie 4.0” in the Making—a Case of Future Told, Tamed, and Traded. *Nanoethics*, 11(1), 107–121.
- Plomp, T., & Nieveen, N. (2013). *Educational Design Research*. Netherlands: Netherlands Institute for Curriculum Development (SLO).
- Sanders, M., & Makotsa, D. (2016). The Possible Influence of Curriculum Statements and Textbooks on Misconceptions: The Case of Evolution. *Education as Change*, 20(1), 216–238.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan RASCH Pada Assessment Pendidikan*. Cimahi: Trim Komunikata Publishing House.
- Sutopo. (2016). Pemahaman Mahasiswa Tentang Konsep Konsep Dasar Gelombang Mekanik. *Jurnal Pendidikan Fisika Indonesia*, 12(2), 41–53.
- Taslidere, E. (2016). Development and Use of a Three-Tier Diagnostic Test to Assess High School Students' Misconceptions about the Photoelectric Effect. *Research in Science and Technological Education*, 34(2), 164–186.
- Uyulgan, M. A., Akkuzu, N., & Alpat, S. (2014). Assessing the Students' Understanding Related to Molecular Geometry Using a Two-Tier Diagnostic Test. *Journal of Baltic Science Education*, 13(6), 839–855.
- Vitharana, P. R. K. A. (2015). Student Misconceptions about Plant Transport a Sri Lankan Example. *European Journal of Science and Mathematics Education*, 3(3), 275–288.
- Vrabec, M., & Proksa, M. (2016). Identifying Misconceptions Related to Chemical Bonding Concepts in the Slovak School System Using the Bonding Representations Inventory as a Diagnostic Tool. *Journal of Chemical Education*, 93(8), 1364–1370.
- Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the Critical Values for Lawshe's Content Validity Ratio. *Measurement and Evaluation in Counseling and Development*, 45(3), 197–210.
- Yang, D. C., & Lin, Y. C. (2015). Assessing 10- to 11-Year-Old Children's Performance and Misconceptions in Number Sense Using a Four-Tier Diagnostic Test. *Educational Research*, 57(4), 368–388.
- Yang, T. C., Chen, S. Y., & Hwang, G. J. (2015). The Influences of a Two-Tier Test Strategy on Student Learning: A Lag Sequential Analysis Approach. *Computers and Education*, 82(C), 366–377.
- Zaleha, Z., Samsudin, A., & Nugraha, M. G. (2017). Pengembangan Instrumen Tes Diagnostik VCCI Bentuk Four-Tier Test pada Konsep Getaran. *Jurnal Pendidikan Fisika dan Keilmuan (JPFK)*, 3(1), 36.
- Zulfikar, A., Samsudin, A., & Saepuzaman, D. (2017). Pengembangan Terbatas Tes Diagnostik Force Concept Inventory Berformat Four-Tier Test. *Jurnal Wahana Pendidikan Fisika*, 2(1), 43–49.