



PENGEMBANGAN INSTRUMEN FOUR-TIER DIAGNOSTIC TEST UNTUK MENDETEKSI MISKONSEPSI SISWA KELAS X PADA MATERI IMPULS

Fungky Iqlima Nasyidiah¹ *, Parsaoran Siahaan¹, Dedi Sasmita¹

Departemen Pendidikan Fisika, Fakultas Pendidikan dan Ilmu Pengetahuan Alam,
Universitas Pendidikan Indonesia, Bandung, 40154 Indonesia

*E-mail: fungkiqlimansydh@student.upi.edu

ABSTRAK

Miskonsepsi merupakan satu dari permasalahan yang dialami siswa pada proses pembelajaran terutama ketika mereka belajar di sekolah. Materi Impuls adalah satu dari konsep abstrak dalam ilmu fisika, lebih tepatnya pada bagian mekanika. Impuls menjadi sulit dipahami ketika diimplementasikan pembelajaran konvensional dan sangat jarang penelitian yang mengungkap miskonsepsi siswa pada materi impuls. Maka dari itu, penelitian ini bertujuan mengembangkan Instrumen Four-Tier Diagnostic Test untuk mengungkap miskonsepsi siswa pada materi impuls. Penelitian ini menggunakan desain penelitian yaitu 3D+1I (defining, designing, developing, dan implementing). Four-Tier Diagnostic Test terdiri dari empat tingkatan, yaitu tingkat pertama berisi pertanyaan dengan empat pilihan jawaban tertutup dan satu pilihan jawaban terbuka, tingkat kedua berisi tingkat keyakinan menjawab pada tier pertama, tingkat ketiga berisi beberapa pilihan alasan tertutup dan satu pilihan alasan terbuka, dan tingkat keempat berisi tingkat keyakinan dalam memilih pilihan alasan. Subjek penelitian ialah siswa kelas X di SMA Kota Bandung. Pengujian validitas Four-Tier Diagnostic Test oleh validator menunjukkan valid. Reliabilitas Instrumen Four-Tier Diagnostic Test menunjukkan konsistensi jawaban dari siswa lemah, namun kualitas butir-butir soal dalam instrument aspek realibilitasnya bagus. Soal tes memiliki 6 soal sukar, 2 soal mudah dan 2 soal sangat mudah. Serta untuk daya pembeda soal tes memiliki 5 soal sangat baik, 2 soal baik, 2 soal kurang baik dan 1 soal jelek. Walau begitu soal tetap dapat diterima karena tidak ada nilai yang menunjukkan nilai negatif. Terdapat 58% siswa terkategori partial negative pada nomor soal 10 dan 13% siswa terkategori miskonsepsi pada nomor soal 2 dan 10.

Kata Kunci: Four-Tier Diagnostic Test , Impuls, Miskonsepsi, dan Model Rasch

ABSTRACT

Misconception is one of the problems experienced by students in the learning process, especially when study at school. Impulse is one of the abstract concepts in physics, especially mechanics section. Impulses become difficult to understand when conventional learning is implemented and very rarely research that reveals students' misconceptions on impulse material. Therefore, this study aims to develop a Four-Tier Diagnostic Test Instrument to reveal students' misconceptions on impulse material. This study used a research design, namely 3D + 1I (defining, designing, developing, and implementing). The Four-Tier Diagnostic Test consists of four levels, namely the first level contains questions with four closed answer options and one open answer option, the second level contains the confidence level of answering in the first tier, the third level contains several closed reasons and one open choice of reasons, and the fourth level contains the level of confidence in choosing the choice of reason. The research subjects were students of class X in SMA Bandung City. The validity test of the Four-Tier Diagnostic Test by the validator shows that it is valid. The reliability of the Four-Tier Diagnostic Test instrument shows the consistency of the answers from students is weak, but the quality of the items in the instrument's reliability aspect is good. The test questions have 6 difficult questions, 2 easy questions and 2 very easy questions. As well as for distinguishing power the test questions have 5 very good questions, 2 good questions, 2 bad questions and 1 bad question. Even so, the problem can still be accepted because there is no value that shows a negative value. There were 58% of students categorized as partial negative in question numbers 10 and 13% of students categorized as misconceptions in question numbers 2 and 10.

Keywords: Four-Tier Diagnostic Test , Impulse, Miconceptions, dan Rasch Model

PENDAHULUAN

Fisika merupakan ilmu pengetahuan yang mempelajari mengenai gejala alam secara umum. Fisika mempelajari tentang materi, energi, dan kejadian alam, baik yang bersifat makroskopik maupun mikroskopik yang berakitan dengan perubahan zat dan energi [4][18]. Mempelajari fisika pada jenjang Sekolah Menengah Atas (SMA) memiliki beberapa tujuan yaitu salah satunya agar siswa menguasai konsep Fisika dengan baik, hal tersebut tercantum pada Permendikbud Nomor 69 tahun 2013. Namun pada prosesnya untuk membuat siswa menguasai konsep ditemukan beberapa kendala salah satunya miskonsepsi, sebab miskonsepsi mampu menurunkan nilai akademis siswa [17]. Konsep merupakan hal yang mendasar, pemahaman yang kokoh dengan keutuhan konsep memainkan peran penting dalam membantu siswa mengembangkan pengetahuan dasar dan struktur pemahaman mereka, serta menerapkan konsep yang tepat mampu meningkatkan kemampuan problem-solving, sehingga siswa dapat mengembangkan keahlian dan kompetensinya [17]. Kebanyakan siswa mengatakan bahwa yang menyebabkan fisika itu sulit untuk dipahami karena konsepnya yang abstrak, rumusnya yang berbelit-belit dan contohnya kurang jelas sehingga menimbulkan miskonsepsi [27].

“ A misconception is not an incorrect answer that comes from a lack of knowledge or from a mistake “ maksudnya miskonsepsi itu bukan jawaban yang berasal dari lack of knowledge atau dari kesalahan [11]. Miskonsepsi dapat berupa konsep awal, kesalahan, hubungan yang tidak benar antara konsep-konsep, gagasan intuitif atau pandangan yang naif [22]. Beberapa ahli memiliki pandangannya tersendiri, menurut Novak miskonsepsi sebagai suatu interpretasi konsep-konsep dalam suatu pernyataan yang tidak dapat diterima [22]. Menurut Feldisine miskonsepsi sebagai suatu kesalahan dan hubungan yang tidak benar antara konsep-konsep [22]. Beberapa kategori miskonsepsi

pada siswa, diantaranya ialah 1)Preconceived misunderstanding merupakan miskonsepsi berdasarkan pengalaman 2)Incomplete or partial understanding merupakan miskonsepsi yang hanya memahami sebagian 3)Wrong interpretations and comprehensions merupakan miskonsepsi berdasarkan kesalahan interpretasi dan pengertian suatu konsep 4)Vernacular misunderstanding [17]. Miskonsepsi digunakan sebagai perbedaan dengan konsep secara ilmiah yang disepakati teorinya oleh para ahli dikarenakan sering dipakai pada literature [13]. Berdasarkan literatur diatas maka dapat disimpulkan bahwa miskonsepsi merupakan pemahaman yang tidak sesuai dengan teori sesungguhnya dari para ahli melainkan berdasarkan pengalaman dan kebiasaan.

Berdasarkan literature kesalahpahaman konsep selalu dikatakan sebagai miskonsepsi, pada umumnya terjadi diantara para siswa terutama siswa baru yang sering kali bermasalah pada performa akademik mereka diberbagai disiplin ilmu, bahkan seringkali siswa mendapatkan hasil performa yang buruk dan kemampuan problem-solving yang buruk pula [17]. Jika performa belajar siswa terus buruk menyebabkan permasalahan lain dari diri siswa itu sendiri. Hasilnya penelitian Aufschneider dan Rogge bahwa siswa yang mengalami miskonsepsi terjadi akibat kebiasaan siswa dalam memahami suatu materi tanpa berdasarkan konsep yang sesungguhnya melainkan berdasarkan pengalamannya dalam kehidupan sehari-hari [7]. Siswa masuk ke dalam kelas tidak dengan pemahaman yang kosong melainkan memiliki preconception yang kuat dan sudah bertahan lama berdasarkan pengalaman yang mereka alami. Dan apabila miskonsepsi terus dibiarkan dimiliki siswa maka dapat menjadi kuat dan sulit untuk dikoreksi.

Miskonsepsi banyak terjadi dalam bidang fisika, hasil penelitian Wandersee, Mintzes dan Noval menjelaskan bahwa hampir seluruh bidang fisika terjadi miskonsepsi, yang dibuktikan dari 700 studi mengenai miskonsepsi bidang fisika terdapat

300 studi mengenai miskonsepsi tentang mekanika; 159 tentang listrik; 70 tentang panas, optik, dan sifat-sifat materi; 35 tentang bumi dan antariksa; serta 10 studi tentang fisika modern [22]. Dalam jurnal yang berjudul *Four-Tier Diagnostic Test To Identify Misconceptions In Geometrical Optics*, mendapatkan hasil bahwa sebanyak 31.37% miskonsepsi level rendah; 52.94% miskonsepsi level sedang; and 15.69% miskonsepsi level tinggi, dari 107 sample kelas X SMA yang telah mendapatkan pembelajaran mengenai materi optic geometri [10].

Salah satu teknik mendiagnosis atau menilai miskonsepsi siswa yaitu dengan tes diagnostic miskonsepsi [15]. Salah satu tes untuk mengidentifikasi miskonsepsi yaitu menggunakan four tier diagnostic test seperti yang telah dilakukan oleh Gurel, Caleon, dan peneliti lainnya, alat tes ini untuk membantu peneliti dalam mengidentifikasi atau menilai miskonsepsi siswa dengan menghadirkan pertanyaan, jawaban serta skala penilaian keyakinan siswa dalam menjawab pertanyaan yang tersedia [9] [13]. Keuntungan dari Four-tier-diagnostic-test yaitu dapat membedakan antara ketiadaan konsep (*lack of knowledge*) dengan miskonsepsi maksudnya terpisah antara main tiers dan reasoning tiers, tidak adanya konsep yang dimiliki biasanya menggunakan three-tier-test namun memiliki nilai eror yang rendah sedangkan miskonsepsi tinggi sedangkan four-tier bebas dari nilai eror dan penampakan kekurangan konsepnya lebih benar [14]. Keunggulan yang dimiliki tes diagnostik empat tingkat adalah guru dapat: (1) membedakan tingkat keyakinan jawaban dan tingkat keyakinan alasan yang dipilih siswa sehingga dapat menggali lebih dalam tentang kekuatan pemahaman konsep siswa, (2) mendiagnosis miskonsepsi yang dialami siswa lebih dalam, (3) menentukan bagian-bagian materi yang memerlukan penekanan lebih, (4) merencanakan pembelajaran yang lebih baik untuk membantu mengurangi miskonsepsi siswa [2] [19].

Maka dari itu pentingnya mengidentifikasi miskonsepsi pada siswa guna menghindari permasalahan miskonsepsi ini karena jika performa belajar siswa terus buruk menyebabkan permasalahan lain dari diri siswa itu sendiri. Maka dari itu penelitian ini sebagai salah satu langkah untuk mengidentifikasi dan mengetahui penyebab miskonsepsi pada siswa.

METODE

Penelitian ini merupakan penelitian kuantitatif deskriptif.

1. Tempat dan waktu pelaksanaan penelitian

Penelitian ini dilakukan menggunakan media evaluasi berbasis web Google Form. Populasi penelitian ini adalah siswa kelas X di SMA PGRI 1 Bandung yang diduga memiliki miskonsepsi berdasarkan hasil observasi. Sampel penelitian ini berjumlah 63 siswa (40 perempuan dan 23 laki-laki) kelas X (MIPA 1 dan MIPA 2). Waktu penelitiannya dimulai dari tanggal 9 Juli 2020 hingga 11 Juli 2020.

2. Model dan tahap pengembangan

Desain penelitian yaitu 3D+1I (*defining, designing, developing, dan implementing*). Desain pada penelitian ini, terdapat empat tahapan, yaitu Tahapan *defining*, mencari literatur penelitian mengenai miskonsepsi pada materi impuls. Tahapan *designing*, merancang instrument tes pada materi impuls dalam bentuk two-tier open-ended test. Tahapan *developing*, mengembangkan instrument tes two-tier open-ended menjadi instrument tes four-tier open-ended melalui proses studi pendahuluan ke siswa dan judgment oleh ahli. Tahap implementasi, menggunakan instrument tes four-tier open-ended untuk diberikan kepada siswa.

2. Pengambilan dan Analisis Data

Metode pengambilan data menggunakan instrument four-tier test yang dilakukan dengan bantuan media evaluasi berbasis web Google-form. Instrument four-tier test sebagai alat untuk mendeteksi miskonsepsi siswa pada materi impuls. Link Google-form disebar kepada siswa melalui aplikasi whatsapp, aplikasi ini dipilih dikarenakan mudah diakses dan hampir semua siswa menggunakannya.

Analisis data yang dilakukan meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan interpretasi instrument four-tier

tes untuk mengungkap miskonsepsi siswa. Uji validitas menggunakan rasio validitas isi (CVR) yang diusung oleh Lawshe [16] dan dikembangkan oleh Wilson [25], uji validitas dilakukan berdasarkan hasil judgment oleh 7 ahli (5 dosen dan 2 guru fisika). Uji Reliabilitas menggunakan analisis Rasch pada aplikasi MINISTEP 4.5.3 dengan output Table 3.1 Summary Statistics. Tingkat kesukaran soal dianalisis dengan Rasch menggunakan aplikasi MINISTEP 4.5.3 dengan output Table 1 Variable (Wright maps) dan output Table 13 Item Measure untuk mendapat gambaran mengenai kekuatan butir soal. Perhitungan daya pembeda digunakan untuk mengukur sejauhmana suatu butir soal mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu [5]. Semakin tinggi indeks diskriminasi daya pembeda maka termasuk soal yang sangat baik, ini berarti bahwa soal tersebut dapat membedakan antara peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi [5]. Untuk menghitung daya pembeda setiap butir soal dapat menggunakan analisis Rasch Model dengan aplikasi WINSTEP 4.5.3 dengan output Table 10 Item Fit Order yang dianalisis dengan analisis Rasch. Teknik analisis untuk mengidentifikasi miskonsepsi menggunakan tabel kriteria konsep siswa yang dikembangkan oleh [3] [12].

HASIL DAN PEMBAHASAN

Defining Instrumen Four-Tier Diagnostic Test

Tahapan defining, mencari literatur penelitian mengenai miskonsepsi pada materi impuls.

Designing Instrumen Four-Tier Diagnostic Test

Tahapan designing, merancang instrument tes pada materi impuls dalam bentuk *two-tier open-ended test*.

Developing Instrumen Four-Tier Diagnostic Test

Tahapan developing, mengembangkan instrument tes *two-tier open-ended* menjadi instrument tes *four-tier open-ended* melalui proses pengambilan opsi alasan dari siswa

dan meminta komentar dari para ahli. Setiap butir soal yang dikembangkan terdiri atas empat tingkatan diantaranya 1) Tingkatan pertama berupa pertanyaan soal dan 4 pilihan jawaban tertutup dan satu pilihan jawaban terbuka, 2) Tingkatan kedua berupa tingkat keyakinan atas jawaban yang dipilih, 3) Tingkatan ketiga berupa beberapa pilihan alasan atas jawaban yang dipilih tertutup dan satu pilihan alasan terbuka, 4) Tingkatan keempat berupa tingkat keyakinan atas alasan yang dipilih. Instrumen yang digunakan merujuk pada Caleon [9]. Soal tes didesain menggunakan model 3D+1I (defining, designing, developing, dan implementing). Secara garis besar hasil dari produk yang dikembangkan disajikan pada Tabel 1 dan Gambaran instrument *four-tier open-ended* akhir yang digunakan disajikan pada Gambar 1.

Tabel 1
Gambaran produk *Four-tier Diagnostic Test*

Produk yang Dikembangkan	Isi
Kisi-kisi Soal <i>Four-tier diagnostic test</i>	Miskonsepsi, konsep ilmiah, indikator soal, soal empat tingkat, jumlah soal, kunci jawaban
Soal <i>four-tier test</i>	1) Tingkatan pertama berupa pertanyaan soal dan 4 pilihan jawaban tertutup dan satu pilihan jawaban terbuka, 2) Tingkatan kedua berupa tingkat keyakinan atas jawaban yang dipilih, 3) Tingkatan ketiga berupa beberapa pilihan alasan atas jawaban yang dipilih tertutup dan satu pilihan alasan terbuka, 4) Tingkatan keempat berupa tingkat keyakinan atas alasan yang dipilih.
Lembar Jawaban	Nama, Nama Sekolah, alamat email, petunjuk pengerjaan soal, kolom jawaban, kolom tingkat keyakinan jawaban, kolom alasan, kolom tingkat keyakinan alasan
Pedoman Penskoran	Pedoman dalam memberikan skor dan menentukan hasil tes
Pedoman Interpretasi Hasil Jawaban Siswa	Pedoman untuk mengklasifikasi kategori konsep siswa

7.1. Sebuah bola Softball bermassa 85 gram dilempar mendatar ke kanan dengan kelajuan 20 m/s. Setelah dipukul, bola bergerak ke kiri dengan kelajuan 20 m/s. Berapakah impulsnya yang diberikan oleh pemukul pada bola ... a. - 0,6 N.s b. - 1,7 N.s c. - 2,1 N.s d. - 3,4 N.s e.												
7.2. Tingkat Keyakinan untuk jawaban yang Anda pilih <table style="width: 100%; text-align: center;"> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> </tr> <tr> <td>Hanya menebak</td> <td>Sangat Tidak Yakin</td> <td>Tidak Yakin</td> <td>Yakin</td> <td>Sangat Yakin</td> <td>Benar Yakin</td> </tr> </table>	1	2	3	4	5	6	Hanya menebak	Sangat Tidak Yakin	Tidak Yakin	Yakin	Sangat Yakin	Benar Yakin
1	2	3	4	5	6							
Hanya menebak	Sangat Tidak Yakin	Tidak Yakin	Yakin	Sangat Yakin	Benar Yakin							
7.3. Alasan yang tepat atas jawaban anda adalah... a. Impulsnya sangat besar, karena momentum yang diberikan tidak seimbang dengan impuls b. Kecepatannya berbanding terbalik dari awalnya positif menjadi negative c. Bola mendapatkan gaya impulsive dari pemain pada selang waktu yang sangat singkat, benda tersebut akan mengalami perubahan momentum d.												
7.4. Tingkat Keyakinan untuk alasan yang Anda pilih <table style="width: 100%; text-align: center;"> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> </tr> <tr> <td>Hanya menebak</td> <td>Sangat Tidak Yakin</td> <td>Tidak Yakin</td> <td>Yakin</td> <td>Sangat Yakin</td> <td>Benar Yakin</td> </tr> </table>	1	2	3	4	5	6	Hanya menebak	Sangat Tidak Yakin	Tidak Yakin	Yakin	Sangat Yakin	Benar Yakin
1	2	3	4	5	6							
Hanya menebak	Sangat Tidak Yakin	Tidak Yakin	Yakin	Sangat Yakin	Benar Yakin							

Gambar 1

Gambaran instrument akhir *four-tier open-ended*

Tahap validitas judgment ahli dilakukan dengan dua cara, yaitu validitas konstruk dan validitas isi. Untuk menguji validitas konstruk dapat digunakan pendapat para ahli (*judgment experts*), setelah instrument dikonstruksi tentang aspek-aspek yang hendak diukur berlandaskan teori maka selanjutnya dikonsultasikan dengan para ahli [6]. Para ahli akan memberi keputusan apakah instrument tersebut dapat digunakan tanpa perbaikan, ada perbaikan, atau dirombak total. Pada penelitian ini menggunakan pendapat dari tujuh judgment ahli/validator, diantaranya ialah 5 dosen ahli dan 2 guru fisika. Proses validasi oleh para ahli dilakukan hingga instrumen *four-tier* dinyatakan layak untuk dijadikan sebagai instrumen penelitian.

Setelah proses revisi selesai langkah selanjutnya ialah menghitung validitas isi, untuk mendukung hasil validasi para ahli dilakukan uji validitas yang digunakan ialah validitas isi melalui pendekatan yang diusulkan oleh Lawshe [16]. Validitas isi merupakan modal dasar dalam suatu instrumen penelitian, sebab kesahihah/validitas isi akan menyatakan keterwakilan aspek yang diukur dalam instrumen, dan dipandang dari segi isi instrumen yang diberikan [26]. Pendekatan validitas isi yang dipilih berdasarkan usulan Lawshe [16] menggunakan rasio validitas isi (CVR) dengan mengukur derajat kesepakatan para ahli dari satu item dan yang dapat mengekspresikan tingkat validitas isi melalui indikator tunggal yang berkisar dari -1 sampai 1. Hal ini relevan dengan yang dilakukan oleh Setiap penilai / *subject matter experts* (SME) yang terdiri dari panel juri (para ahli) menjawab pertanyaan untuk setiap *item* dengan tiga pilihan jawaban yaitu (1) valid tanpa revisi, (2) valid dengan revisi, (3) tidak valid. Menurut Lawshe, jika lebih dari setengah panelis menunjukkan bahwa item penting/esensial, maka *item* tersebut

memiliki setidaknya validitas isi. Maka validitas untuk setiap item instrument *Four-tier* berdasarkan sembilan indikator penilaian dapat dilihat pada tabel 2.

Tabel 3. Hasil Indeks CVR Validasi Ahli

No. Soal	Validator							Indeks CVR	Keputusan
	I	II	III	IV	V	VI	VII		
1.	2,78	3	3	3	2,78	2,67	3	0,968	Digunakan
2.	2,78	3	3	3	2	2,56	3	0,904	Digunakan
3.	2,78	2,89	3	3	1	2,67	3	0,683	Digunakan
4.	2,78	2,89	3	3	1,67	2,44	3	0,841	Digunakan
5.	2,78	2,78	3	3	2	2,56	3	0,905	Digunakan
6.	2,78	3	3	3	2,11	2,67	3	0,968	Digunakan
7.	2,67	3	3	3	1,78	2,56	3	0,841	Digunakan
8.	2,67	3	3	3	1,89	2,56	3	0,873	Digunakan
9.	2,78	3	3	3	3	2,56	3	0,937	Digunakan
10.	2,44	2,89	3	3	3	2,67	3	0,937	Digunakan

Implementing Instrumen Four-Tier Diagnostic Test

Tahap implementasi, menggunakan instrument tes *four-tier open-ended* untuk diberikan kepada siswa. Kemudian setelah mendapatkan hasil jawaban siswa, dilakukan analisis mengenai nilai validitas butir soal, reliabilitas, tingkat kesukaran, daya pembeda dan mengungkap miskonsepsi siswa berdasarkan kombinasi jawaban siswa.

Uji Validitas

Uji validitas instrument *four-tier* menggunakan software MINISTEP 4.5.3 dengan output Table 10 item (column): fit order yang dianalisis dengan analisis Rasch. Pengukuran uji validitas instrument dilihat berdasarkan nilai logarithm odd unit (logit) pada outfit mean square (MNSQ), outfit Z-standard (ZSTD), dan point-measure correlation (PT-MEASURE CORR).

Hasil uji validitas instrument *four-tier* dapat dilihat pada Gambar 2

Fungky Iqlima Nasyyidah dkk, Pengembangan Instrumen *Four-Tier Diagnostic Test* Untuk Mendeteksi Miskonsepsi Siswa Kelas X Pada Materi Impuls

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MODEL MEASURE	S.E.	INFIIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-AL CORR.	EXP.	EXACT OBS% MATCH	Item	
8	14	45	.44	.35	1.12	.90	1.68	2.52	.16	.34	70.5	70.5	E8
3	26	45	-.87	.33	1.19	1.47	1.29	1.69	.23	.42	56.8	67.0	E3
2	11	45	.82	.37	1.08	.50	1.06	.29	.24	.31	75.0	75.4	E2
9	20	45	-.23	.33	1.08	.74	1.06	.45	.32	.39	63.6	66.7	E9
1	34	44	-1.98	.40	.99	.02	.96	.00	.44	.43	81.4	80.8	E1
10	11	42	.67	.37	.99	-.01	.91	-.25	.34	.32	73.2	73.9	E10
5	17	45	.10	.33	.89	-.93	.82	-.96	.48	.37	72.7	67.8	E5
6	20	45	-.23	.33	.88	-1.05	.85	-.97	.50	.39	72.7	66.7	E6
4	14	45	.44	.35	.87	-.90	.80	-.87	.47	.34	75.0	70.5	E4
7	11	45	.82	.37	.84	-.90	.74	-.86	.47	.31	79.5	75.4	E7
MEAN	17.8	44.6	.00	.35	.99	.02	1.02	.11			72.0	71.5	
P.SD	7.1	.9	.83	.02			.27	1.11			6.8	4.5	

Gambar 2. OUTFIT (MNSQ) dan (ZSTD) setiap butir soal

Selanjutnya, Nilai OUTFIT (MNSQ) dan (ZSTD) dimasukan pada kriteria nilai yang diterima, untuk memeriksa kesesuaian butir soal. Berdasarkan kriteria, maka validitas untuk setiap item instrument four-tier dapat dilihat pada Tabel 4.

Tabel 4 Validitas item instrument *four-tier*

No. Soal	Nilai outfit (MNSQ)	Nilai outfit (ZSTD)	Keterangan
1.	0,99	0,02	Diterima
2.	1,08	0,50	Diterima
3.	1,19	1,47	Diterima
4.	0,87	-0,90	Diterima
5.	0,89	-0,93	Diterima
6.	0,88	-1,05	Diterima
7.	0,84	-0,90	Diterima
8.	1,12	0,90	Diterima
9.	1,08	0,74	Diterima
10.	0,99	-0,1	Diterima

Pada tabel 4 terlihat bahwa seluruh item soal menunjukkan kesesuaian dengan syarat kriteria, maka dapat dikatakan bahwa seluruh item soal mempunyai kecenderungan fit yang artinya seluruh item soal berfungsi normal melakukan pengukuran dan tidak terindikasi terjadinya miskonsepsi.

Realibilitas Butir Soal Instrumen Four-tier Diagnostic Test

Uji reliabilitas pada Rasch memiliki cara penafsiran berbeda dengan reliabilitas dalam teori klasik yang memiliki harga tunggal, sedangkan uji reliabilitas pada Rasch mampu menghasilkan pengukuran berbeda antara individu berkemampuan tinggi dan individu

berkemampuan rendah [21]. Analisis Rasch dapat menampilkan beberapa nilai realibilitas, diantaranya adalah person reability, item reability, dan Cronbach alpha. Person reability menunjukkan konsistensi jawaban peserta didik, item reability menunjukkan kualitas item tes, sementara Cronbach alpha menunjukkan nilai interaksi Antara person dan item dari instrument four-tier secara keseluruhan. Hasil uji reabilitas butir soal dapat dilihat pada Gambar 2.

SUMMARY OF 10 MEASURED (NON-EXTREME) Item

	TOTAL SCORE	COUNT	MODEL MEASURE	S.E.	INFIIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	17.8	44.6	.00	.35	.99	-.02	1.02	.11
SEM	2.4	.3	.28	.01	.04	.29	.09	.38
P.SD	7.1	.9	.83	.02	.11	.86	.27	1.13
S.SD	7.5	1.0	.88	.02	.12	.90	.28	1.19
MAX.	34.0	45.0	.82	.40	1.19	1.47	1.68	2.52
MIN.	11.0	42.0	-1.98	.33	.84	-1.05	.74	-.97
REAL RMSE	.36	TRUE SD	.75	SEPARATION	2.08	Item	RELIABILITY	.81
MODEL RMSE	.35	TRUE SD	.76	SEPARATION	2.14	Item	RELIABILITY	.82
S.E. OF Item MEAN	= .28							

SUMMARY OF 44 MEASURED (NON-EXTREME) Person

	TOTAL SCORE	COUNT	MODEL MEASURE	S.E.	INFIIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	4.0	9.9	-.46	.75	.99	.03	1.01	.06
SEM	.2	.1	.14	.02	.04	.14	.06	.13
P.SD	1.6	.4	.89	.11	.25	.89	.39	.82
S.SD	1.6	.4	.90	.11	.25	.90	.39	.83
MAX.	7.0	10.0	.99	1.11	1.56	2.45	2.01	2.01
MIN.	1.0	8.0	-2.50	.68	.56	-2.00	.24	-1.55
REAL RMSE	.80	TRUE SD	.38	SEPARATION	.48	Person	RELIABILITY	.19
MODEL RMSE	.76	TRUE SD	.46	SEPARATION	.60	Person	RELIABILITY	.27
S.E. OF Person MEAN	= .14							

SUMMARY OF 45 MEASURED (EXTREME AND NON-EXTREME) Person

	TOTAL SCORE	COUNT	MODEL MEASURE	S.E.	INFIIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	4.0	9.9	-.53	.78				
SEM	.3	.1	.15	.03				
P.SD	1.7	.4	1.01	.20				
S.SD	1.7	.4	1.02	.20				
MAX.	7.0	10.0	.99	1.88				
MIN.	.0	8.0	-3.86	.68				
REAL RMSE	.84	TRUE SD	.56	SEPARATION	.67	Person	RELIABILITY	.31
MODEL RMSE	.80	TRUE SD	.61	SEPARATION	.77	Person	RELIABILITY	.37
S.E. OF Person MEAN	= .15							

Person RAW SCORE TO MEASURE CORRELATION = .87
 CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .30 SEM = 1.43

Gambar 2. Nilai person reability, item reability, dan Cronbach alpha.

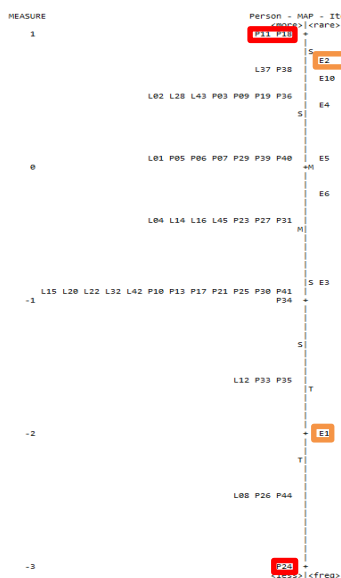
Berdasarkan analisis Rasch maka nilai person reability untuk pengukuran non-ekstrim (menghilangkan satu data yang tergolong ekstrim) yaitu 0,19 dan nilai person reability untuk pengukuran gabungan antara ekstrim dengan non-ekstrim yaitu 0,31 keduanya termasuk kategori lemah serta nilai

item reability 0,81 termasuk kategori bagus. Dapat disimpulkan bahwa konsistensi jawaban dari siswa lemah masuk dalam kategori lemah, namun kualitas butir-butir soal dalam instrument aspek realibilitasnya bagus. Nilai Cronbach alpha dari instrument four-tier ialah 0,30 sehingga termasuk kategori buruk, yang artinya bahwa interaksi antara person dan butir-butir soal secara keseluruhan buruk.

Karakteristik Butir Soal Four-Tier Diagnostic Test

Soal tes dikatakan baik apabila valid, reliabel, serta memiliki tingkat kesukaran dan daya pembeda yang baik. Pada bagian ini akan dibahas mengenai karakteristik butir soal dilihat dari tingkat kesukaran dan daya pembeda dari setiap butir soal pada Four-tier Diagnostic Test.

Tingkat kesukaran soal digunakan untuk mengukur seberapa besar derajat kesukaran setiap butir soal, jika suatu soal memiliki tingkat kesukaran seimbang, maka dikatakan bahwa soal tersebut baik [5]. Tingkat Kesukaran dapat dianalisis dengan Rasch menggunakan aplikasi MINISTEP 4.5.3 dengan output Table 1 Variable (Wright maps) dan output Table 13 Item Measure untuk mendapat gambaran mengenai kekuatan butir soal. Adapun hasil dapat dilihat pada Gambar 3 mengenai Variable (Wright maps).



Gambar 3. Variabel (Wright) maps instrument four-tier

Kode E diikuti oleh nomor pada bagian kanan Gambar 3.3 menunjukkan nomor butir soal dan kode Huruf diikuti oleh nomor pada bagian kiri menunjukkan nomor urut siswa. Gambar 3 merupakan suatu peta yang menggambarkan sebaran kemampuan siswa dalam menjawab soal dan sebaran tingkat kesukaran soal, kode soal E2 dan E7 menunjukkan kesukaran butir soal yang paling tinggi dapat dijawab oleh 20 siswa dengan kode L01, L02, L04, P05, P06, P09, P10, P11, L15, P18, P19, L22, P25, P29, P36, L37, P38, P41, L43 dan L45 dan kode soal E1 menunjukkan kesukaran butir soal yang paling rendah dapat dijawab oleh 34 siswa dengan kode L01, P03, L04, P05, P06, P07, L08, P09, P10, P11, L12, P13, L14, L15, P17, P18, P19, L20, L22, P23, P25, L28, P29, P30, P31, P35, P36, L37, P38, P39, P40, L42, L43, dan L45. Secara lebih rinci, tingkat kesukaran instrument four-tier pada Gambar 3.4 dapat ditentukan berdasarkan nilai logit output Table 13 Item Measure pada software MINISTEP.

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT MATCH 085%	MATCH EXP%	Item
2	11	45	.82	.37	1.08	.50	1.06	-.29	.24	.31	75.0	75.4	E2
7	11	45	.82	.37	.84	-.90	.74	-.86	.47	.31	79.5	75.4	E7
10	11	42	.67	.37	.99	-.01	.91	-.25	.34	.32	73.2	73.9	E10
4	14	45	.44	.35	.87	-.90	.80	-.87	.47	.34	75.0	70.5	E4
8	14	45	.44	.35	1.12	.90	1.68	2.52	.16	.34	70.5	70.5	E8
5	17	45	.10	.33	.89	-.93	.82	-.96	.48	.37	72.7	67.8	E5
6	20	45	-.23	.33	.88	-1.05	.85	-.97	.50	.39	72.7	66.7	E6
9	20	45	-.23	.33	1.08	.74	1.06	.45	.32	.39	63.6	66.7	E9
3	26	45	-.87	.33	1.19	1.47	1.29	1.69	.23	.42	56.8	67.0	E3
1	34	44	-1.98	.40	.99	.02	.96	.00	.44	.43	81.4	80.8	E1
MEAN	17.8	44.6	.00	.35	.99	.0	1.02	.1			72.0	71.5	
P.SD	7.1	.9	.83	.02	.11	.9	.27	1.1			6.8	4.5	

Gambar 4. Nilai MEASURE setiap butir soal dan Standar Deviasi (SD)

Data tingkat kesukaran instrument four-tier pada Gambar 4 diurutkan berdasarkan nilai logit dari yang terbesar yaitu kode soal E2 dan E7 sampai nilai logit dari yang terkecil yaitu kode soal E1. Nilai logit terbesar menunjukkan tingkat soal yang tinggi, hal ini berkorespondensi dengan kolom total score, yaitu menyatakan berapa jumlah jawaban benar. Adapun pada item measure terdapat informasi nilai standar deviasi 0,83. Jika nilai ini dikombinasikan dengan rata-rata logit

maka dapat dikelompok berdasarkan tingkat kesulitannya. 0,0 logit+1SD adalah satu kelompok soal sulit, lebih besar dari +1SD adalah soal yang sangat sukar; 0,0 logit-1SD adalah satu kelompok soal mudah, lebih besar dari -1SD adalah soal yang sangat mudah [21]. Artinya ada empat kategori tingkat kesukaran soal yaitu, 1) Sangat sulit 2) Sulit 3) Mudah 4) Sangat Mudah.

Tabel 5.

Hasil tingkat kesukaran tiap butir soal

No. Soal	Tingkat Kesukaran (TK)	Interpretasi
1.	-1,98	Sangat Mudah
2.	0,82	Sukar
3.	-0,87	Sangat Mudah
4.	0,44	Sukar
5.	0,10	Sukar
6.	-0,23	Mudah
7.	0,82	Sukar
8.	0,44	Sukar
9.	-0,23	Mudah
10.	0,67	Sukar

Tabel 5 menunjukkan bahwa instrument four-tier memiliki tingkat kesukaran yang cukup beragam, dimulai dari soal sangat mudah, mudah, dan sulit. Sehingga sebarannya tingkat kesukarannya dapat dikatakan merata.

Daya pembeda digunakan untuk mengukur sejauhmana suatu butir soal mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu [5]. Untuk menghitung daya pembeda setiap butir soal dapat menggunakan analisis Rasch Model dengan aplikasi WINSTEP 4.5.3 dengan output Table 10 Item Fit Order yang dianalisis dengan analisis Rasch. Adapun nilai PTMEASURE-AL COOR untuk instrument four-tier dapat dilihat pada Gambar 5

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MODEL MEASURE	MODEL S.E.	INFIT		OUTFIT		PTMEASURE-AL		EXACT EXP.	MATCH OBS%	Titem
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.			
8	14	45	.44	.35	1.12	.90	1.68	2.52	A .16	.34	70.5	70.5	E8
3	26	45	-.87	.33	1.19	1.47	1.29	1.69	B .23	.42	56.8	67.0	E3
2	11	45	.82	.37	1.08	.50	1.06	.29	C .24	.31	75.0	75.4	E2
9	20	45	-.23	.33	1.08	.74	1.06	.45	D .32	.39	63.6	66.7	E9
1	34	44	-1.98	.40	.99	.02	.96	.00	E .44	.43	81.4	80.8	E1
10	11	42	.67	.37	.99	-.01	.91	-.25	e .34	.32	73.2	73.9	E10
5	17	45	.10	.33	.89	-.93	.82	-.96	d .48	.37	72.7	67.8	E5
6	20	45	-.23	.33	.88	-1.05	.85	-.97	c .50	.39	72.7	66.7	E6
4	14	45	.44	.35	.87	-.90	.80	-.87	b .47	.34	75.0	70.5	E4
7	11	45	.82	.37	.84	-.90	.74	-.86	a .47	.31	79.5	75.4	E7
MEAN	17.8	44.6	.00	.35	.99	.0	1.02	.1			72.0	71.5	
P.SD	7.1	.9	.83	.02	.11	.9	.27	1.1			6.8	4.5	

Gambar 5. Nilai PTMEASURE-AL COOR setiap butir soal

Berdasarkan nilai PTMEASURE-AL COOR pada gambar 5, maka daya pembeda untuk tiap soal dapat dilihat pada Tabel 6

Tabel 6

Hasil PTMEASURE-AL COOR untuk tiap butir soal

No. Soal	PTMEASURE-AL COOR	Interpretasi
1.	0,44	Sangat Baik
2.	0,24	Kurang Baik
3.	0,23	Kurang Baik
4.	0,47	Sangat Baik
5.	0,48	Sangat Baik
6.	0,50	Sangat Baik
7.	0,47	Sangat Baik
8.	0,16	Jelek
9.	0,32	Baik
10.	0,34	Baik

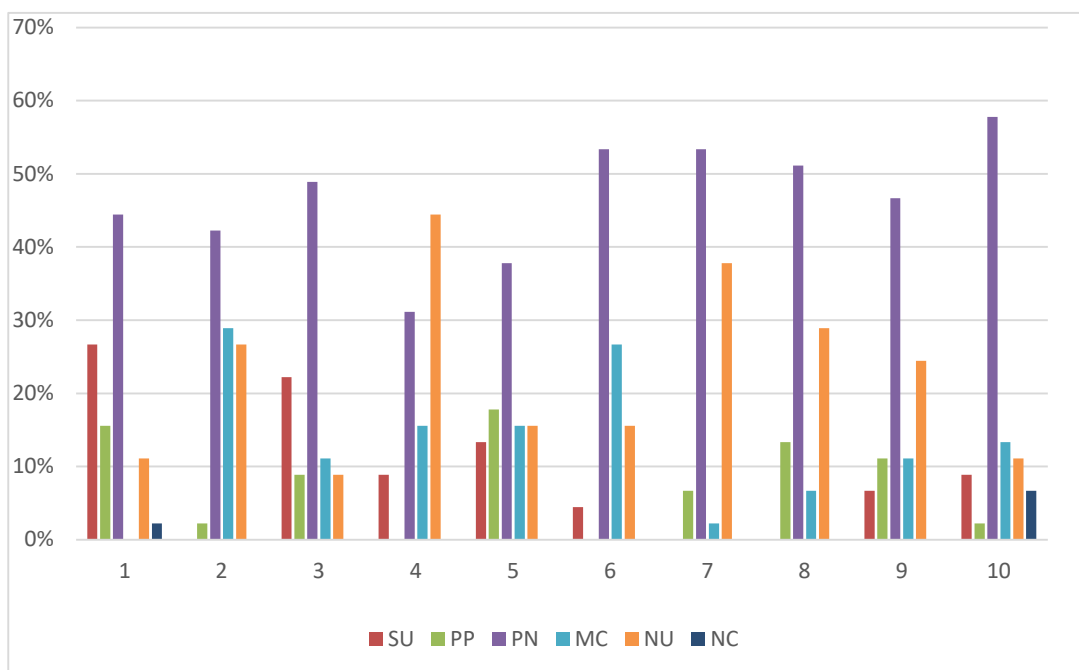
Tabel 6 menunjukkan bahwa tidak terdapat daya pembeda yang bernilai negative sehingga seluruh soal dapat diterima.

Identifikasi Miskonsepsi Siswa

Berdasarkan Tabel yang dikembangkan Gurel dan Aminudin, dkk serta dengan hasil jawaban yang diberikan siswa memberikan suatu gambaran bahwa pemahaman siswa mengenai konsep impuls beraneka ragam, baik antar siswa maupun antar item soal [3] [12]. Hasilnya dapat dihitung presentase dari setiap kategori konsepsi (SU, PP, PN, MC, NU, NC), hasil perhitungan diinterpretasikan dalam bentuk grafik pada Gambar 5.

Hasilnya akan difokuskan pada miskonsepsi dan partial negative yang dialami oleh siswa, bahwa kategori Partial Negative (PN) atau Partial Understanding with Specific

Misconception dengan siswa terbanyak yang terkategori partial negative disetiap item soal, hal ini dapat dilihat pada gambar 7.



Gambar 7
Grafik Presentase Kategori Konsepsi Siswa

Presentase Partial Negative (PN) tertinggi berada pada soal nomor 10 sebanyak 26 siswa (58%) dengan kode L01, P05, P05, P06, P07, P08, P10, P11, P13, L16, L20, P21, L22, P23, P25, P26, P27, P29, P30, P32, P33, P39, L42, L43, P44, dan L45 serta terendah pada soal nomor 4 sebanyak 14 siswa (31%) dengan kode L01, P05, P06, P07, P011, L12, L14, L15, L16, P21, P34, P36, P40, dan L43. Kategori Misconception (MC) tertinggi berada pada soal nomor 2 sebanyak 13 siswa (29%) dengan kode L02, P07, P09, P13, P18, P23, P27, dan L37 serta terendah pada soal nomor 1 sebanyak nol siswa (0%). Kategori No Understanding (NU) tertinggi berada pada soal nomor 4 sebanyak 20 siswa (44%) dengan kode P03, L08, P10, P13, P17, P19, P20, P22, P23, P24, P27, P29, P31, P32, P33, P35, P39, P41, P44, dan L45 serta terendah pada soal nomor 3 sebanyak 4 siswa (9%) dengan kode P07, L08, P23, dan L43. Hasil menunjukkan bahwa hampir seluruh siswa mengalami miskonsepsi, baik itu sebagian maupun secara spesifik dari materi impuls yang diujikan. Siswa memiliki kepercayaan

yang sama pada suatu konsep yang tidak sesuai dengan konsep ahli. Hal ini perlu diwaspadai melihat miskonsepsi yang terjadi telah meluas dan akan menjadi masalah bagi siswa dalam menerima pengetahuan baru.

KESIMPULAN

Instrumen four-tier diagnostic test yang dihasilkan terdiri dari kisi-kisi soal tes, petunjuk pengerjaan soal, soal tes, kunci jawaban, lembar jawaban menggunakan google-form, pedoman penskoran, dan pedoman interpretasi hasil. Soal tes memiliki empat tingkatan, diantaranya 1) Tingkatan pertama berupa pertanyaan soal dan pilihan jawaban, 2) Tingkatan kedua berupa tingkat keyakinan atas jawaban yang dipilih, 3) Tingkatan ketiga berupa pilihan alasan atas jawaban yang dipilih, 4) Tingkatan keempat berupa tingkat keyakinan atas alasan yang dipilih. Produk akhir yang dihasilkan berjumlah 10 butir, terdiri dari 9 indikator dan 3 sub pokok bahasan. Semua validator yang

berjumlah 7 orang menyatakan tiap butir soal tes valid sehingga dapat diujikan kepada subjek penelitian. Reliabilitas Instrumen Four-Tier Diagnostic Test menunjukkan konsistensi jawaban dari siswa lemah, namun kualitas butir-butir soal dalam instrument aspek realibilitasnya bagus. Soal tes memiliki 6 soal sukar, 2 soal mudah dan 2 soal sangat mudah. Serta untuk daya pembeda soal tes memiliki 5 soal sangat baik, 2 soal baik, 2 soal kurang baik dan 1 soal jelek. Walau begitu soal tetap dapat diterima karena tidak ada nilai yang menunjukkan nilai negatif. Terdapat 58% siswa terkategori partial negative pada nomor soal 10 dan 13% siswa terkategori miskonsepsi pada nomor soal 2 dan 10.

REFERENSI

- [1] Arikunto, S. (2013). *Prosedur Penelitian Suatu Pendekatan Praktik* (15th ed.). Jakarta: Rineka Cipta.
- [2] Cahyo, A. N. (2013). *Panduan Aplikasi Teori-Teori Belajar Mengajar Teraktual dan Terpopuler*. Jogjakarta: Diva Press.
- [3] Darmadi, H. (2013). *Metode Penelitian Pendidikan*. Bandung: Alfabeta.
- [4] Hartati, B. (2010). Pengembangan Alat Peraga Gaya Gesek untuk Meningkatkan Kemampuan Berfikir Kritis Siswa SMA. *Jurnal Pendidikan Fisika Indonesia*, 1-5.
- [5] Hastuti, T. W. (2014). Kemampuan Berfikir Kritis Siswa SMA Muhammadiyah 2 Surakarta pada Pembelajaran Biologi Berbasis Praktikum. 1-15.
- [6] Husnindar, Ikhsan, M., & Rizal, S. (2014). Penerapan Model Pembelajaran Berbasis Masalah untuk Meningkatkan Kemampuan Berpikir Kritis dan Disposisi Matematis Siswa. *Jurnal Didaktik Matematika*, 1(1), 71-82.
- [7] Mubarok, C., & Sulisty, E. (2014). Penerapan Model Pembelajaran Discovery Learning Terhadap Hasil Belajar Siswa Kelas X TAV Pada Standar Kompetensi Melakukan Instalasi Sound System Di SMK Negeri 2 Surabaya. *Jurnal Pendidikan Teknik Elektro*, 215-221.
- [8] Nurlaila, N., Suparmi, & Sunarno, W. (2013). Pembelajaran Fisika Dengan PBL Menggunakan Problem Solving dan Problem Posing Ditinjau dari Kreativitas Dan Ketrampilan Berpikir Kritis Siswa. *JURNAL INKUIRI*, 2(2), 114-123.
- [9] Putu S, N. L., Gede A, A. A., & Dantes, N. (2015). Pengaruh Model Pembelajaran Berbasis Masalah (PBL) terhadap Hasil Belajar, Ketrampilan Berfikir Kritis dan Sikap Ilmiah dalam Pembelajaran Fisika Kelas XI IPA SMA N Kuta tahun Pelajaran 2014/2015. 1-7.
- [10] Sudargo, F., & Asiah, S. (2010). Kemampuan Pedagogik Calon Guru dalam Meningkatkan Kemampuan Berfikir Kritis dan Ketrampilan Proses Siswa melalui Pembelajaran Berbasis Praktikum. *Jurna; Pengajaran MIPA*, 4-12.
- [11] Sudargo, F., & Asiah, S. (2010). Kemampuan Pedagogik Calon Guru dalam Meningkatkan Kemampuan Berfikir Kritis dan Ketrampilan Proses Siswa Melalui Pembelajaran Berbasis Praktikum. 1-9.
- [12] Sukardi. (2007). *Metodologi Penelitian Pendidikan*. Jakarta: PT Bumi Aksara.
- [13] Widiadnyana, Sadia, & Suastra. (2014). Pengaruh Model Discovery Learning Terhadap Pemahaman Konsep IPA dan Sikap Ilmiah Siswa SMP. *e-Journal Program Pascasarjana Universitas Pendidikan Ganesha*.