



Development of a Four Tier Diagnostic Test Instrument Based on iSpring Suite 9 to Identify Understanding of Concepts in Students

Pengembangan Instrumen Four Tier Diagnostic Test Berbasis iSpring Suite 9 Untuk Mengidentifikasi Pemahaman Konsep Pada Mahasiswa

Asri Nurhafsari¹, Febri Rismaningsih^{2*}

Universitas Islam Syekh-Yusuf, Indonesia

Email: anurhafsari@unis.ac.id ; frismaningsih@unis.ac.id

Abstract

Evaluation is important to measure the achievement of learning outcomes. Multiple choices, form of test, is commonly used in learning evaluation. However, in the real situation shows that multiple choice tests make students choose the answers randomly. This study was conducted to provide a solution by developing a Four Tier Diagnostic Test based on iSpring Suite 9 to identify students' conceptual understanding of Static Fluid. This research methodology employed Research & Development (R&D). The subject in this study were students of the Faculty of Engineering at Universitas Islam Syekh-Yusuf who are studying Basic Physics. Data collection techniques use observation, interviews, tests, validation sheets, and questionnaires. The validation results of material experts are 93%, media experts are 84% so that they are included in the very feasible category. The validity of the test showed that 35 items were valid, and 7 questions were invalid. The reliability of the test is 0.93. The index of difficulty consists of 3 easy, 35 moderate and 4 difficult. The discriminating power of each item included 16 very good questions, 9 good questions, 11 enough questions, 5 bad questions and 1 very bad item. Student response to this instrument is 74.83% included in the high category. This test can identify four levels of students' understanding of concepts, namely understanding concepts, not understanding concepts, partially understanding and misconceptions. Based on the results, the Four Tier Diagnostic Test based on iSpring Suite 9 has met the criteria for a good instrument and is suitable for use.

Keywords: development; four tier diagnostic test; understanding of concepts; iSpring Suite 9

Abstrak

Evaluasi penting dilakukan untuk mengetahui ketercapaian hasil belajar. Salah satu bentuk tes yang sering digunakan dalam evaluasi pembelajaran adalah bentuk multiple choice. Namun, kenyataan yang dijumpai saat ini bentuk multiple choice membuat mahasiswa hanya asal memilih jawaban. Penelitian ini bertujuan untuk memberikan solusi dengan mengembangkan instrumen *Four Tier Diagnostic Test* berbasis iSpring Suite 9 untuk mengidentifikasi pemahaman konsep mahasiswa pada materi Fluida Statis. Penelitian ini merupakan jenis *Research & Development* (R&D). Subjek penelitian ini adalah mahasiswa Fakultas Teknik Universitas Islam Syekh-Yusuf yang sedang mengambil mata kuliah Fisika Dasar. Pengumpulan datanya menggunakan pengamatan, wawancara, instrumen tes, instrumen validasi, dan angket. Hasil penilaian validasi ahli materi 93%, ahli media 84% sehingga termasuk dalam kategori sangat layak. Validitas tes menunjukkan 35 butir soal valid dan 7 soal tidak valid. Reliabilitas tesnya adalah 0,93. Indeks kesukaran pada instrumen ini yaitu 3 soal mudah, 35 soal sedang dan 4 soal sukar. Daya pembeda setiap butir soal antara lain 16 soal sangat baik, 9 butir soal baik, 11 soal cukup baik, 5 soal jelek dan ada 1 soal sangat jelek. Respon mahasiswa terhadap instrumen ini adalah 74,83% termasuk dalam kategori tinggi. Hasil tes ini dapat mengidentifikasi empat tingkatan pemahaman konsep mahasiswa yaitu paham konsep, tidak paham konsep, paham sebagian dan miskonsepsi. Berdasarkan hasil yang diperoleh, instrumen *Four Tier Diagnostic Test* berbasis *iSpring Suite 9* ini telah memenuhi kriteria instrumen yang baik dan layak digunakan.

Kata Kunci: pengembangan; *four tier diagnostic test*; pemahaman konsep; *iSpring Suite 9*

DOI	:	doi.org/10.24903/bej.v5i2.1345
Received	:	July 2023
Accepted	:	July 2023
Published	:	August 2023
Copyright and License	:	Authors retain copyright and grant the journal right of first publication with the work simultaneously licensed under a Creative Commons Attribution 4.0 International License that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal. 

1. PENDAHULUAN

Hasil belajar merupakan bagian terpenting dalam suatu proses pembelajaran. Ketika mahasiswa telah melalui serangkaian proses pembelajaran diharapkan mampu mengubah perilaku kearah yang lebih baik dari sebelumnya. Namun pada kenyataannya, seringkali mahasiswa tidak mencapai tujuan belajarnya dan tidak menunjukkan perubahan perilaku menjadi lebih baik. Hal ini mengindikasikan bahwa mahasiswa mengalami adanya kesulitan belajar dalam mencapai target hasil belajar, terutama pemahaman konsep yang masih rendah. Pemahaman konsep dapat diartikan sebagai pengetahuan yang dipelajari oleh mahasiswa terkait topik maupun gagasan tertentu dengan cara yang bermakna dan terintegrasi (Taslidere, E., Yıldırım, 2023). Konsep yang dipahami terkadang tidak sesuai dengan pemahaman yang telah dikonstruksi para ahli sebelumnya (Rismaningsih & Nurhafsari, 2022). Untuk dapat menggali adanya pemahaman konsep secara mendalam pada mahasiswa maka perlu adanya suatu evaluasi pembelajaran.

Evaluasi tidak dapat terlepas dari kegiatan belajar mengajar, melalui evaluasi seorang pendidik dapat memperoleh informasi terkait pencapaian hasil belajar yang telah dilakukan. Evaluasi merupakan proses akhir dalam suatu pembelajaran yang bertujuan untuk mengidentifikasi hasil belajar dan mengukur pemahaman konsep materi yang telah dipelajari. Ada dua cara yang dapat dipakai dalam melaksanakan evaluasi pembelajaran yaitu dengan menggunakan teknik tes dan nontes (Hayati et al., 2020). Teknik tes merupakan teknik untuk menilai hasil belajar mahasiswa dengan memberikan pertanyaan-pertanyaan terkait materi yang telah dipelajari. Salah satu bentuk tes yang sering dipakai oleh pendidik saat melakukan evaluasi pembelajaran yaitu bentuk tes pilihan ganda (*multiple choice*).

Sebenarnya tidak ada yang keliru dengan bentuk tes *multiple choice*. Namun, realita yang ada saat ini tes yang berbentuk *multiple choice* menjadikan mahasiswa hanya asal dalam

menentukan jawaban (Rovita, 2020). Hal ini juga terjadi pada pelaksanaan evaluasi mata kuliah Fisika Dasar, salah satu materi yang dikaji yaitu Fluida Statis. Pada evaluasi materi Fluida Statis membutuhkan alasan maupun proses pengerjaan untuk menentukan jawaban yang tepat. Saat mengerjakan evaluasi, sebagian besar mahasiswa hanya memilih jawaban tanpa melakukan proses pengerjaan sehingga tidak mengetahui alasan mengapa mereka memilih jawaban tersebut. Pelaksanaan evaluasi pembelajaran yang demikian menjadikan pendidik tidak dapat menggali informasi yang lebih mendalam mengenai jawaban mahasiswa. Oleh karena itu, pendidik memerlukan bentuk tes lain yang dapat digunakan dalam evaluasi dengan tujuan mendapatkan informasi yang lebih detail mengenai pemahaman konsep dan juga hasil belajar mahasiswa. Alternatif solusi yang dapat dipakai untuk mengatasi permasalahan ini adalah dengan melakukan evaluasi yang menggunakan *Four Tier Diagnostic Test (FTDT)*. Hal ini sejalan dengan penelitian yang dilakukan oleh (Fariyani et al., 2015) bahwa *Four Tier Diagnostic Test* mampu memberikan gambaran terkait kemampuan siswa dalam materi Optik Geometri.

Four Tier Diagnostic Test mempunyai keunggulan dibandingkan tes diagnostik yang telah ada sebelumnya antara lain: a) Dapat membedakan tingkat keyakinan jawaban mahasiswa dan tingkat keyakinan alasan yang dipilih mahasiswa sehingga dapat menggali lebih dalam tentang kekuatan pemahaman mahasiswa. b) Mendiagnosis miskonsepsi yang dialami oleh mahasiswa secara mendalam. c) Mengidentifikasi bagian-bagian materi yang memerlukan penekanan lebih, dan d) Merancang pembelajaran yang lebih baik untuk membantu mengurangi miskonsepsi mahasiswa (Kaltakci-Gurel et al., 2017).

Pasca pandemi kegiatan belajar mengajar secara bertahap sudah mulai dilaksanakan secara tatap muka. Evaluasi yang dilakukan pun banyak yang kembali seperti semula yaitu *paper-based test*. Hal ini tentunya perlu adanya suatu adaptasi baik dari pendidik maupun mahasiswa. Berdasarkan hasil wawancara dengan beberapa mahasiswa S1 Fakultas Teknik Universitas Islam Syekh-Yusuf (UNIS) menyatakan bahwa meskipun pembelajaran sudah bertahap dilakukan secara tatap muka, namun dalam pemberian tes dapat diselingi dengan menggunakan media yang lebih bervariasi agar mahasiswa lebih termotivasi, tertarik dan lebih mudah dalam memahami konsep khususnya pada konsep materi Fluida Statis.

Salah satu program yang dapat digunakan oleh pendidik dalam melaksanakan evaluasi pembelajaran yang kreatif dan inovatif adalah *iSpring Suite 9*. Program *iSpring Suite 9* merupakan aplikasi komputer untuk membuat evaluasi pembelajaran yang menarik, efektif, efisien dan mempunyai kemampuan yang lebih unggul karena menyediakan berbagai pilihan

bentuk soal tes evaluasi seperti *short answer, matching, true or false*, dan lain sebagainya (Ninawati et al., 2021)(Wardhono et al., 2019). Selain mudah dalam penggunaannya, *iSpring Suite 9* juga mudah dalam pembuatannya karena merupakan aplikasi tambahan untuk Powerpoint sehingga lebih familiar terhadap menu-menunya. Dengan menggunakan program ini, pendidik dapat membuat berbagai bentuk soal tes yang sesuai kebutuhan untuk evaluasi pembelajaran. Output dari aplikasi ini ada dalam format html atau web, sehingga ke depannya dapat diintegrasikan ke dalam aplikasi yang telah dimodifikasi untuk dapat digunakan di komputer maupun *handphone* dengan bantuan aplikasi lain yaitu web 2 apk. Produk akhirnya berupa aplikasi yang dapat digunakan pada komputer dan *handphone* dengan cara membagikannya melalui link bit.ly sehingga soal siap untuk dikerjakan mahasiswa.

Oleh karena itu, peneliti bermaksud mengidentifikasi pemahaman konsep mahasiswa pada materi Fluida Statis menggunakan instrumen *Four Tier Diagnostic Test* berbasis *iSpring Suite 9*. Hasil dari penelitian ini diharapkan mampu menghasilkan alat evaluasi yang efektif untuk mengidentifikasi pemahaman konsep mahasiswa Fakultas Teknik UNIS pada jenjang sarjana. Selain itu dapat memudahkan pendidik dalam mengoreksi jawaban serta memberikan penilaian karena hasil jawaban dalam *iSpring Suite 9* dapat mengidentifikasi benar salahnya maupun yakin atau tidaknya alasan yang diberikan mahasiswa.

2. METODOLOGI

Metode penelitian ini adalah penelitian dan pengembangan atau yang sering disebut *Research and Development (R&D)*. Penelitian ini akan menghasilkan produk berupa alat evaluasi *Four Tier Diagnostic Test* menggunakan *iSpring Suite 9* untuk mengidentifikasi pemahaman konsep yang lebih mendalam pada mahasiswa dalam mata kuliah Fisika Dasar. Tahapan penelitian *R&D* yang dikembangkan oleh (Sugiyono, 2013) ada 10 yaitu : Potensi dan Masalah, Pengumpulan Data, Desain Produk, Validasi Desain, Revisi Desain, Uji Coba Produk, Revisi Produk, Uji Coba Pemakaian, Revisi Produk, dan Produksi Massal. Dari sepuluh langkah tersebut, tahapan yang digunakan dalam penelitian ini disesuaikan dengan kondisi dan kebutuhan saat penelitian sehingga dalam penelitian ini mengadaptasi delapan tahapan saja.

Populasi penelitian ini adalah seluruh mahasiswa S1 Fakultas Teknik Universitas Islam Syekh-Yusuf yang sedang menempuh perkuliahan Fisika Dasar. Sampel dalam penelitian ini diambil melalui teknik *puspositive sampling* yaitu pemilihan sampel yang didasarkan pada pertimbangan peneliti dan dosen pengampu yang bersangkutan. Uji coba instrumen dilakukan

kepada 29 mahasiswa tingkat atas yang telah mengambil matakuliah Fisika Dasar sedangkan uji coba pemakaian dilakukan pada mahasiswa sampel sebanyak 47 orang.

Metode pengumpulan data yang dilakukan dalam penelitian ini adalah wawancara, pengamatan, instrumen tes, instrumen validasi, dan angket. Teknik analisis data yang dilakukan antara lain analisis validasi ahli, analisis instrumen tes, analisis deskriptif dari data tes dan angket penggunaan *Four Tier Diagnostic Test* berbasis *iSpring Suite 9*. Validasi instrumen tes dilakukan oleh ahli materi dan juga ahli media dengan mengisi lembar validasi yang telah disediakan peneliti. Lembar validasi ahli berupa angket yang menggunakan skala Likert dengan nilai 1 (Sangat Kurang), 2 (Kurang), 3 (Cukup), 4 (Baik) dan 5 (Sangat Baik). Hasil validasi ahli kemudian dianalisis untuk mengetahui kualitas instrumen tes diagnostik. Caranya adalah menggunakan rumus sebagai berikut (Sugiyono, 2013).

$$P = \frac{f}{N} \times 100\%$$

Keterangan:

P = persentase skor, f = jumlah skor yang diperoleh dan N = jumlah skor maksimum

Selanjutnya nilai tersebut diinterpretasikan dalam tabel berikut:

Tabel 1. Range Persentase Interpretasi Kelayakan

Penilaian	Kriteria
$80\% < x \leq 100\%$	Sangat Layak
$60\% < x \leq 80\%$	Layak
$40\% < x \leq 60\%$	Cukup Layak
$20\% < x \leq 40\%$	Tidak Layak
$0\% < x \leq 20\%$	Sangat Tidak Layak

Sumber: (Khoir et al., 2020)

Selain divalidasi oleh ahli, instrumen tes diolah dan dianalisis juga dengan bantuan *software* ANATES untuk mengetahui hasil perhitungan validitas, reliabilitas, indeks kesukaran dan daya pembeda pada setiap butir soal. Selanjutnya, analisis data angket respon mahasiswa terhadap penggunaan *Four Tier Diagnostic Test* berbasis *iSpring Suite 9* dihitung menggunakan persentase secara keseluruhan maupun analisis setiap indikatornya.

3. HASIL

Penelitian ini menghasilkan produk berupa instrumen *Four Tier Diagnostic Test* berbasis *iSpring Suite 9*. Berikut ini adalah rincian dari hasil penelitian yang dilakukan.

3.1 Penyusunan Instrumen *Four-Tier Diagnostic Test*

Tes diagnostik yang dikembangkan ini dilakukan untuk mengidentifikasi pemahaman konsep mahasiswa pada materi Fluida Statis dengan subkonsep yaitu Tekanan Hidrostatik, Hukum Pascal, Hukum Archimedes, Tegangan Permukaan, Kapilaritas dan Viskositas. Adapun interpretasi jawaban yang diberikan responden menurut (Gurel et al., 2015) pada *Four Tier Diagnostic Test* ditunjukkan pada tabel berikut.

Tabel 2. Interpretasi Responden Pada *Four-Tier Diagnostic Test*

No.	Kategori	Kombinasi Jawaban			
		Pilihan (tier I)	Tingkat yakin pertama (tier II)	Alasan (tier III)	Tingkat yakin kedua (tier IV)
1.	Paham Konsep	Benar	Yakin	Benar	Yakin
		Benar	Yakin	Benar	Tidak Yakin
		Benar	Tidak Yakin	Benar	Yakin
		Benar	Tidak Yakin	Benar	Tidak Yakin
2.	Paham Sebagian	Benar	Yakin	Salah	Yakin
		Benar	Yakin	Salah	Tidak Yakin
		Benar	Tidak Yakin	Salah	Yakin
		Benar	Tidak Yakin	Salah	Tidak Yakin
		Salah	Yakin	Benar	Yakin
		Salah	Yakin	Benar	Tidak Yakin
		Salah	Tidak Yakin	Benar	Yakin
		Salah	Tidak Yakin	Benar	Tidak Yakin
3.	Miskonsepsi	Salah	Yakin	Salah	Yakin
4.	Tidak Paham Konsep	Salah	Yakin	Salah	Tidak Yakin
		Salah	Tidak Yakin	Salah	Yakin
		Salah	Tidak Yakin	Salah	Tidak Yakin

Sumber: (Gurel et al., 2015)

Tahapan penyusunan instrumen penelitian *Four Tier Diagnostic Test* ada tiga tahap, yaitu:

a. Tahap Persiapan

Pada tahap ini, peneliti menentukan isi materi, dan menganalisis konsep materi yang dipilih. Selanjutnya peneliti menggali informasi mengenai pemahaman konsep mahasiswa. Informasi ini berasal dari tinjauan literatur dengan mengacu pada penelitian sebelumnya. Setelah mengkaji literatur, dibuat indikator pertanyaan sesuai dengan kompetensi inti dan tujuh kompetensi dasar konsep Fluida Statis, dan kemudian membuat kisi-kisi pertanyaan wawancara klinikal mahasiswa.

b. Tahap Pembuatan Instrumen Tingkat Pertama (Tier I)

Pada tahapan ini dilakukan wawancara klinikal dengan menggunakan pertanyaan yang telah disiapkan sebelumnya. Wawancara klinikal dilaksanakan sebagai referensi untuk pembentuk tier pertama. Hasil wawancara klinikal selanjutnya dianalisis dan dimasukkan sebagai pertanyaan tingkat pertama (tier 1).

c. Tahap Pembuatan Instrumen Tier Alasan

Pada tahapan ini dibuat rancangan instrumennya yaitu 42 soal formatif dan terdapat 5 pilihan kemungkinan jawaban, serta ditambahkan soal alasan untuk mahasiswa isi sesuai jawaban mahasiswa yang dihasilkan dari tahap pembuatan instrumen tingkat pertama. Tes ini mirip dengan tes *multiple choice* dengan alasan terbuka. Alasan yang diberikan mahasiswa selanjutnya akan dianalisis oleh peneliti dan dikembangkan berdasarkan hasil jawaban mahasiswa dan hasil telaah studi kepustakaan. Selanjutnya instrumen dua tingkat tersebut dikembangkan dengan menambahkan 2 tingkatan keyakinan dan diletakkan setelah soal pertanyaan dan setelah soal alasan.

Hasil instrumen tes yang telah dikembangkan dapat terlihat pada tabel berikut ini.

Tabel 3. Produk *Four Tier Diagnostic Test*

Produk	Isi
Kisi-kisi Uji Instrumen Tes Diagnostik (awal)	Sub pokok bahasan, indikator soal, tingkat kognitif, nomor soal dan jumlah soal 42
Soal Uji coba <i>Four Tier Diagnostic Test</i>	42 soal tes fluida statis, pilihan jawaban, tingkat keyakinan terhadap jawaban, pilihan alasan dan tingkat keyakinan alasan.
Kisi-kisi Uji Instrumen Tes Diagnostik (revisi)	Sub pokok bahasan, indikator soal, tingkat kognitif, nomor soal dan jumlah soal 35
Soal revisi <i>Four Tier Diagnostic Test</i> (Uji Pemakaian)	35 soal-soal tes fluida statis, pilihan jawaban, tingkat keyakinan jawaban, pilihan alasan dan tingkat keyakinan alasan.
Soal <i>Four Tier Diagnostic Test</i> berbasis <i>iSpring Suite 9</i>	Nama, NIM, Prodi, email, No.Hp, Petunjuk pengerjaan soal, soal/pertanyaan tes fluida statis, pilihan jawaban, tingkat keyakinan jawaban, pilihan alasan dan tingkat keyakinan alasan.
Kunci Jawaban <i>Four Tier Diagnostic Test</i>	Nomor soal, pilihan jawaban benar dan alasan yang benar
Pedoman penskoran <i>Four Tier Diagnostic Test</i>	Pedoman dalam memberikan penilaian/skor dan menentukan hasil tes
Pedoman Interpretasi Hasil <i>Four Tier Diagnostic Test</i>	Pedoman untuk mengetahui klasifikasi jawaban yang diberikan oleh mahasiswa

Penyebaran butir soal berdasarkan tingkat kognitif, dapat dilihat pada tabel berikut ini.

Tabel 4. Penyebaran Butir Soal

Level ranah kognitif	Instrumen Awal		Instrumen revisi	
	No Soal	Banyaknya	No Soal	Banyaknya
C1	23	1	-	-
C2	1,4,11,17,27,31,38	7	1, 8, 14, 22, 26, 32	6
C3	8,13,19,29,30	5	5, 10, 16, 24, 25	5
C4	2,3,5,6,10,12,15,16,18, 24,28,33,34,35,36,37, 39,40,41,42	20	2,3,4,7,9,12,13,15, 19,23,28,29,30,31, 33,34,35	17
C5	9,14,20,21,25,32	6	6,11,17,18,20,27	6
C6	7,22,26	3	21	1
	Jumlah Soal	42	Jumlah Soal	35

3.2 Validitas Four Tier Diagnostic Test

Instrumen tes diagnostik yang telah disusun selanjutnya akan diuji kelayakannya berdasarkan penilaian oleh validator. Validasi instrumen dilaksanakan oleh ahli materi dan juga ahli media. Penilaian yang dilakukan oleh ahli materi bertujuan untuk menilai kelayakan isi instrumen tersebut. Hasil validasi ahli materi terhadap instrumen tes diagnostik mendapatkan penilaian sebesar 93%. Hal ini berarti kelayakan instrumen tes masuk dalam kategori “Sangat Layak” untuk digunakan. Validasi ahli yang lain dilakukan oleh ahli media yang bertujuan untuk menilai kelayakan *Four Tier Diagnostic Test* menggunakan media *iSpring Suite 9*. Hasil penilaian validasi ahli media mendapatkan nilai sebesar 84%. Hal ini berarti kelayakan instrumen termasuk ke dalam kategori “Sangat Layak” dan perlu adanya perbaikan. Perbaikan yang dilakukan sesuai saran ahli media terkait pemilihan jenis dan ukuran huruf, kesesuaian proporsi warna dan tampilan *background*.

Selanjutnya soal tes yang telah divalidasi ahli kemudian diujikan terhadap mahasiswa peserta uji instrumen yaitu sebanyak 29 mahasiswa. Uji coba ini dilaksanakan untuk mengetahui detail setiap butir soal yang digunakan sehingga instrumen tes diagnostik ini dapat mengukur miskonsepsi mahasiswa. Pada awal uji coba instrumen diberikan soal sebanyak 42 soal terkait Fluida Statis dengan berbagai tingkat kognitif. Hasil uji coba ini dianalisis dengan bantuan ANATES dan mendapatkan hasil bahwa dari 42 soal yang diberikan terdapat 7 butir (16,7 %) soal yang tidak valid, dan 35 soal (83,3 %) yang dinyatakan valid. Soal-soal yang tidak valid adalah soal nomor 3, 4, 7, 22, 23, 34, dan 42. Berdasarkan pertimbangan, 7 soal yang tidak valid dibuang dan yang digunakan dalam pengembangan *Four Tier Diagnostic Test* adalah sebanyak 35 soal dimana penomoran soalnya disesuaikan kembali.

3.3. Reliabilitas Four Tier Diagnostic Test

Reliabilitas adalah konsistensi soal dalam melakukan apa yang seharusnya dinilai. Pengujian reliabilitas juga dilakukan dengan menggunakan software ANATES. Dari hasil perhitungan terlihat bahwa reliabilitas tes diagnostik ini adalah sebesar 0,93. Hal ini berarti instrumen tes diagnostik yang diujicobakan adalah instrumen yang reliabel.

3.4 Daya Pembeda dan Tingkat Kesukaran Four Tier Diagnostic Test

Instrumen yang baik hendaknya selain harus valid dan reliabel, instrumen juga harus mempunyai tingkat kesukaran dan juga daya pembeda yang baik. Tingkat kesukaran dan daya pembeda tes diagnostik ini dihitung menggunakan software ANATES. Dari 42 soal yang diujicobakan hasilnya adalah 3 butir soal mudah, 35 soal kategori sedang dan 4 soal termasuk sukar. Hal ini terlihat bahwa sebanyak 83,3 % soal yang diberikan kepada mahasiswa merupakan soal dengan tingkat kesukaran sedang. Sedangkan untuk daya pembeda setiap butir soal, rinciannya adalah 16 butir soal mempunyai kategori sangat baik, 9 butir soal mempunyai daya pembeda baik, 11 butir soal cukup baik, 5 butir soal berkategori jelek dan ada 1 butir soal yang daya pembedanya sangat jelek.

3.5 Rekapitulasi Hasil Uji Coba

Berikut ini adalah rekapitulasi dari hasil uji coba instrumen yang telah dilaksanakan tes uji coba pada 29 mahasiswa.

```
REKAP ANALISIS BUTIR
=====
Rata2= 21,00
Simpang Baku= 9,95
KorelasiXY= 0,88
Reliabilitas Tes= 0,93
Butir Soal= 42
Jumlah Subyek= 29

Btr Baru  Btr Asli  D.Pembeda(%)  T. Kesukaran  Korelasi  Sign. Korelasi
1          1          12,50         Sedang        0,279     Signifikan
2          2          62,50         Mudah         0,618     Sangat Signifikan
3          3          0,00          Mudah         -0,196    -
4          4          37,50         Sedang        0,272     -
5          5          75,00         Sedang        0,468     Sangat Signifikan
6          6          62,50         Sedang        0,616     Sangat Signifikan
7          7          37,50         Sedang        0,251     -
8          8          37,50         Sedang        0,327     Signifikan
9          9          75,00         Sedang        0,624     Sangat Signifikan
10         10         75,00         Sedang        0,617     Sangat Signifikan
11         11         62,50         Sedang        0,473     Sangat Signifikan
12         12         37,50         Sedang        0,381     Sangat Signifikan
13         13         100,00        Sedang        0,851     Sangat Signifikan
14         14         75,00         Sedang        0,675     Sangat Signifikan
15         15         100,00        Sedang        0,950     Sangat Signifikan
16         16         50,00         Sedang        0,515     Sangat Signifikan
17         17         62,50         Sedang        0,529     Sangat Signifikan
18         18         37,50         Sedang        0,414     Sangat Signifikan
19         19         12,50         Sedang        0,320     Signifikan
20         20         25,00         Sedang        0,376     Sangat Signifikan
21         21         62,50         Mudah         0,557     Sangat Signifikan
22         22         25,00         Sedang        0,138     -
23         23         37,50         Sedang        0,227     -
24         24         87,50         Sedang        0,833     Sangat Signifikan
25         25         75,00         Sedang        0,653     Sangat Signifikan
26         26         75,00         Sedang        0,712     Sangat Signifikan
27         27         87,50         Sukar         0,552     Sangat Signifikan
28         28         37,50         Sedang        0,356     Sangat Signifikan
29         29         75,00         Sedang        0,518     Sangat Signifikan
30         30         37,50         Sedang        0,356     Sangat Signifikan
31         31         12,50         Sedang        0,284     Signifikan
32         32         87,50         Sedang        0,645     Sangat Signifikan
33         33         87,50         Sedang        0,836     Sangat Signifikan
34         34         12,50         Sukar         0,214     -
35         35         75,00         Sedang        0,572     Sangat Signifikan
36         36         62,50         Sedang        0,508     Sangat Signifikan
37         37         50,00         Sedang        0,645     Sangat Signifikan
38         38         75,00         Sedang        0,445     Sangat Signifikan
39         39         75,00         Sedang        0,756     Sangat Signifikan
40         40         37,50         Sukar         0,445     Sangat Signifikan
41         41         50,00         Sedang        0,496     Sangat Signifikan
42         42         12,50         Sukar         0,196     -
```

Gambar 1. Hasil Rekapitulasi Analisis Butir dengan Menggunakan ANATES

3.6 Angket Respon Mahasiswa

Data ini merupakan data kualitatif yang diambil dari penyebaran kuisisioner kepada mahasiswa mengenai respon mahasiswa terhadap penggunaan *iSpring Suite 9*. Angket ini disebar kepada seluruh mahasiswa yang menjadi subjek penelitian yaitu sebanyak 47 mahasiswa. Angket diisi ketika mahasiswa telah selesai mengerjakan *Four Tier Diagnostic Test* sehingga mahasiswa dapat memberikan tanggapan terkait instrumen ini.

Angket ini berjumlah 19 butir pertanyaan yang terdiri dari indikator kepuasan, kemudahan, efektivitas, efisiensi, kenyamanan, kinerja aplikasi, tampilan aplikasi, ketertarikan, kelengkapan. Hasil skor yang diperoleh sebesar 3341 dari skor total yaitu 4465 didapatkan perhitungan persentasenya yaitu 74,83%. Berdasarkan kualifikasi analisis kriteria kuesioner termasuk dalam kategori tinggi. Oleh karena itu, dapat disimpulkan bahwa Sebagian besar mahasiswa memberikan respon positif terhadap *Four Tier Diagnostic Test* berbasis *iSpring Suite 9*.

3.7 Four Tier Diagnostic Test berbasis iSpring Suite9

Butir-butir soal tes diagnostik yang telah diuji coba dan valid kemudian dibuat ke dalam aplikasi *iSpring Suite 9* sehingga link pengerjaannya dapat disebarakan kepada mahasiswa. Soal yang disusun ke dalam *iSpring Suite* terdiri dari 35 butir soal pertanyaan, 35 butir tingkat keyakinan jawaban, 35 butir soal alasan dan 35 butir tingkat keyakinan alasan yang diberikan. Berikut ini akan diuraikan tampilan dari hasil *Four Tier Diagnostic Test* di *iSpring Suite 9*.

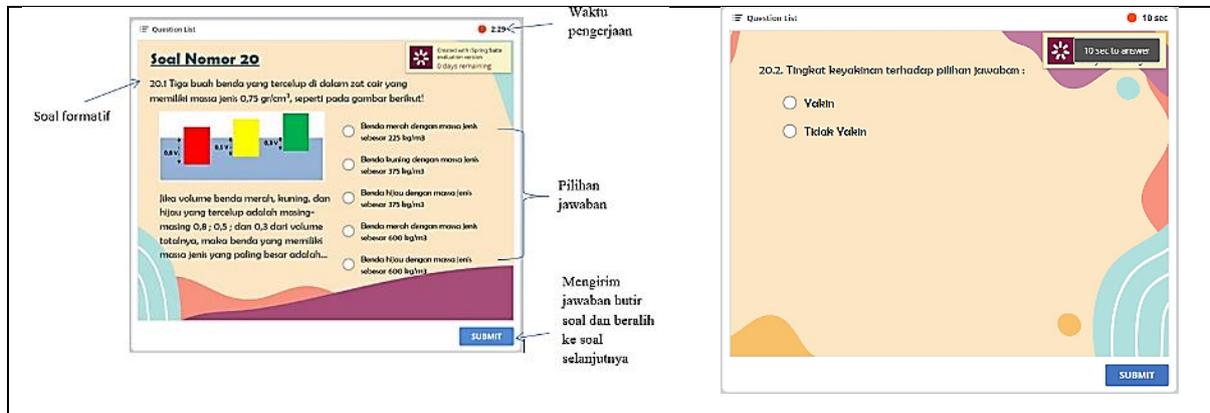
Pada tampilan awal tes diagnostik di *iSpring Suite 9* terdiri dari halaman depan judul tes, identitas mahasiswa yang wajib diisi, dan petunjuk untuk mengerjakan soal. Gambar dari tampilan di *iSpring Suite 9* dapat di lihat pada gambar berikut.



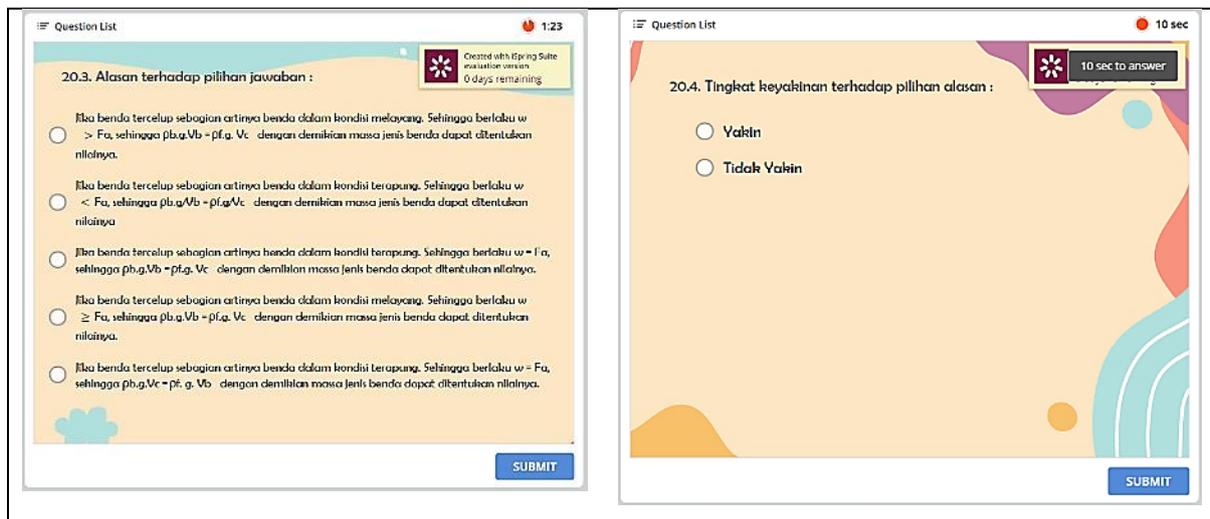
Gambar 2. Tampilan Awal Tes Diagnostik, Halaman Identitas Mahasiswa, Petunjuk Pengisian Soal

Setelah mahasiswa mengisi identitas dan membaca petunjuk soal, maka slide selanjutnya adalah soal-soal tes diagnostik yang terdiri dari 35 butir soal, dimana setiap soal terdapat 4 pertanyaan yang berupa soal formatif, tingkat keyakinan dari soal formatif, soal alasan, dan tingkat keyakinan dari soal alasan. Untuk soal formatif dan soal alasan terdapat 5

pilihan jawaban dan untuk tingkat keyakinan terdapat dua pilihan jawaban. Setiap soal yang ada diberikan waktu untuk mengerjakan, waktu disesuaikan dengan tingkat kesukaran soal (1-2,5 menit), sementara untuk tingkat keyakinan diberikan waktu untuk mengerjakan selama 10 detik. Berikut ini diberikan contoh tampilan soal dari *iSpring Suite 9*.



Gambar 3. Contoh Tampilan Soal Formatif dan Tingkat Keyakinan Jawaban



Gambar 4. Contoh Tampilan Soal Alasan dan Tingkat Keyakinan Soal Alasan

Setelah mahasiswa mengerjakan semua butir soal, maka jawaban akan disubmit. Respon mahasiswa akan terekam dan dikirim kepada email pembuat soal. Selanjutnya setelah selesai submit mahasiswa dapat keluar dari tes diagnostik di *iSpring Suite 9*. Pada akhir tes diagnostik di *iSpring Suite 9* sengaja tidak ditampilkan hasil tes untuk menghindari penyalahgunaan jawaban dari soal tes diagnostik. Hal ini disebabkan oleh waktu tes yang diberikan kepada mahasiswa tidak dapat dilaksanakan secara bersamaan melainkan disesuaikan dengan jadwal perkuliahan yang ada.

3.8. Hasil Four Tier Diagnostic Test Terhadap Pemahaman Konsep Mahasiswa

Hasil *Four Tier Diagnostic Test* selanjutnya dilakukan uji coba pemakaian tes kepada 47 mahasiswa yang menempuh mata kuliah Fisika Dasar untuk menyelesaikan soal tersebut, sehingga peneliti mendapatkan data terkait pemahaman konsep mahasiswa. Jawaban

mahasiswa dikelompokkan ke dalam tingkatan pemahaman yang disesuaikan dengan pola jawaban, yaitu paham konsep, paham sebagian, miskonsepsi, dan tidak paham konsep. Selanjutnya dari hasil pengolahan data tes tersebut, diperoleh persentase pemahaman mahasiswa sesuai tingkatan kategori secara keseluruhan. Gambaran lebih jelas mengenai data tersebut dapat dilihat melalui diagram berikut ini.

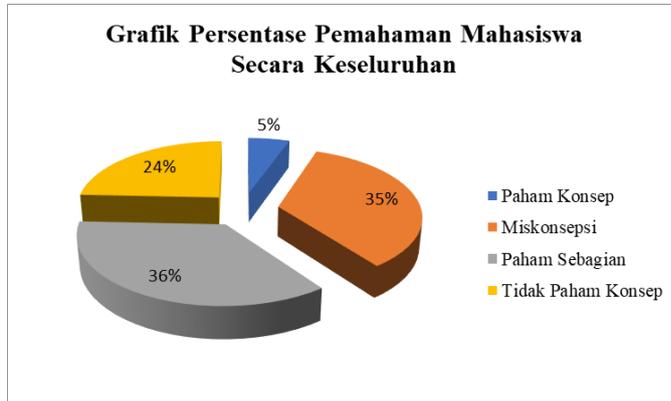


Chart 1 Persentase Pemahaman Mahasiswa Secara Keseluruhan

Gambar 5 menunjukkan persentase pemahaman mahasiswa secara keseluruhan pada tes Fluida Statis. Sebanyak 36 % mahasiswa masuk kategori paham sebagian, 35 % mahasiswa mengalami miskonsepsi, sebanyak 24 % mahasiswa tidak paham konsep materi Fluida Statis, dan persentase paling rendah yaitu hanya 5% mahasiswa yang telah paham konsep materi Fluida Statis.

4. PEMBAHASAN

4.1 Instrumen Four Tier Diagnostic Test

Dalam pengembangan *Four Tier Diagnostic Test* telah dilakukan validasi ahli dan uji coba instrumen. Validasi dilakukan oleh ahli materi dan juga ahli media. Tujuan dilakukannya validasi ahli materi yaitu untuk menilai derajat keabsahan dari materi yang dikembangkan, yakni materi Fluida Statis sedangkan validasi media bertujuan untuk mengukur tingkat kelayakan media yang dikembangkan sebelum digunakan pada tahap pengembangan selanjutnya (Surahman & Surjono, 2017).

Hasil uji coba instrumen didapatkan 35 soal yang valid dan reliabilitas tesnya tinggi. Hal ini sejalan dengan pendapat dari (Saftari & Fajriah, 2019) (Yunita et al., 2017) bahwa instrumen evaluasi yang dibuat harus mempunyai karakteristik instrumen yang baik, sehingga manfaat dari suatu penilaian dapat tercapai. Karakteristik instrumen yang baik diantaranya adalah valid dan reliabel. Pada instrumen *Four Tier Diagnostic Test* yang dikembangkan, sebagian besar soal uji instrumen memiliki tingkat kesukaran sedang. Hal ini sejalan dengan pendapat (Ruseffendi, 2010) bahwa soal-soal yang paling baik yaitu soal yang memiliki tingkat kesukaran sedang sehingga dapat memberikan informasi terkait perbedaan individu. (Arifin,

2013) juga menyebutkan suatu soal sebaiknya mempunyai tingkat kesukaran yang tidak terlalu sukar dan tidak terlalu mudah.

Selain itu instrumen ini mempunyai daya pembeda yang baik, artinya soal-soal yang diberikan mampu menunjukkan perbedaan mahasiswa yang menguasai konsep dengan mahasiswa yang tidak menguasai konsep. Hal ini sesuai dengan pendapat (Ruseffendi, 2010) (Arifin, 2013) bahwa semakin tinggi nilai koefisien daya pembeda maka semakin kuat soal tersebut dapat membedakan antara peserta didik yang pandai dengan peserta didik lemah.

4.2 Angket Respon Mahasiswa

Dari hasil penelitian, responden yang telah mengisi kuesioner sebanyak 47 mahasiswa dengan 19 butir pertanyaan. Hasil perhitungan diperoleh skor 3341 dari skor totalnya 4465 sehingga diperoleh persentase sebesar 74,83%. Berdasarkan kualifikasi analisis kriteria kuesioner termasuk dalam kategori tinggi. Sehingga dapat disimpulkan bahwa mahasiswa memberikan respon positif terhadap *Four Tier Diagnostic Test* berbasis *iSpring Suite 9*. Hal ini sejalan dengan pendapat (Dasmu et al., 2020) bahwa media pembelajaran interaktif berbasis *iSpring Suite 9* dapat digunakan dengan lebih mudah. Selain itu, media pembelajaran *iSpring Suite 9* dalam bentuk *digital content* sesuai dengan kondisi mahasiswa yang merupakan bagian dari generasi Z (Tani & Ekawati, 2019). Dalam penggunaannya *iSpring Suite 9* ini berisikan media gambar, teks, audio, video, serta animasi sehingga sesuai untuk diterapkan sebagai media pemberian soal atau kuis dalam mata kuliah Fisika Dasar baik berupa pilihan ganda atau essay kepada mahasiswa. Hal ini dikarenakan penyajian soal dapat terlihat jelas, lebih menarik dan lebih mudah dipahami.

4.3 Pemahaman Konsep Mahasiswa

Pembahasan hasil uji coba pemakaian *Four Tier Diagnostic Test* berbasis *iSpring Suite 9* dibatasi pada identifikasi pemahaman konsep mahasiswa. Dari hasil penelitian terlihat bahwa *Four Tier Diagnostic Test* dapat mengidentifikasi tingkatan pemahaman konsep mahasiswa yaitu paham konsep, tidak paham konsep, paham sebagian dan miskonsepsi. Adanya tes ini menunjukkan bahwa pendidik tidak hanya melihat mahasiswa paham atau tidak paham konsep saja tetapi lebih mendalam lagi terdapat mahasiswa yang hanya paham sebagian konsep bahkan miskonsepsi. Hal ini senada dengan (Gurel et al., 2017) bahwa *Four Tier Diagnostic Test* mampu membedakan tingkat keyakinan jawaban dan tingkat keyakinan alasan yang dipilih mahasiswa sehingga dapat menggali lebih dalam tentang kekuatan pemahaman mahasiswa. Dengan tes ini pendidik dapat mengetahui seberapa jauh pemahaman konsep mahasiswa sehingga dapat dijadikan refleksi untuk meningkatkan serta memperbaiki metode yang digunakan dalam pembelajaran.

5. KESIMPULAN

Berdasarkan hasil dan pembahasan penelitian, dapat disimpulkan bahwa hasil validasi ahli terhadap instrumen yang dikembangkan ini termasuk dalam kategori sangat layak. Instrumen *Four Tier Diagnostic Test* memiliki validitas dan reliabilitas yang tinggi begitupula tingkat kesukaran dan daya pembeda pada tes ini. Oleh karena itu, instrumen *Four Tier Diagnostic Test* ini telah memenuhi kriteria instrumen yang baik dan layak digunakan. Respon mahasiswa dilihat dari hasil kuesionernya termasuk dalam kategori tinggi, artinya mahasiswa memberikan respon positif terhadap *Four Tier Diagnostic Test* berbasis *iSpring Suite 9*. Tes ini dapat mengidentifikasi tingkatan pemahaman konsep mahasiswa yaitu paham konsep, tidak paham konsep, paham sebagian dan miskonsepsi. Selain itu bagi pendidik, tes ini dapat mengetahui seberapa jauh pemahaman konsep mahasiswa sehingga dapat dijadikan refleksi untuk meningkatkan serta memperbaiki metode yang digunakan dalam pembelajaran. Saran untuk pengembangan selanjutnya, *Four Tier Diagnostic Test* ini dapat menggunakan aplikasi *portable* lainnya dengan menyisipkan animasi maupun audiovisual agar menarik dan tidak bosan saat mengerjakan soal.

6. DAFTAR PUSTAKA

- Arifin, Z. (2013). *Evaluasi Pembelajaran Prinsip, Teknik, Prosedur*. Remaja Rosdakarya.
- Dasmo, Lestari, A. P., & Alamsyah, M. (2020). Peningkatan Hasil Belajar Fisika Melalui Penerapan Media Pembelajaran Interaktif Berbasis Ispring Suite 9. *Prosiding Seminar Nasional Sains*, 1(1). <http://proceeding.unindra.ac.id/index.php/sinasis/article/view/3979>
- Fariyani, Q., Rusilowati, A., & Sugianto. (2015). Pengembangan Four-Tier Diagnostic Test Untuk Mengungkap Miskonsepsi Fisika Siswa SMA Kelas X. *JISE*, 4(2). <http://journal.unnes.ac.id/sju/index.php/jise>
- 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). <https://doi.org/10.12973/eurasia.2015.1369a>
- 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. <https://doi.org/10.1080/02635143.2017.1310094>
- Hayati, U., Ediyani, M., Maimun, M., Anwar, K., Fauzi, M. B., & Suryati, S. (2020). Test Technique as a Tool for Evaluation of Learning Outcomes. *Budapest International Research and Critics Institute (BIRCI-Journal): Humanities and Social Sciences*, 3(2), 1198–1205. <https://doi.org/10.33258/birci.v3i2.961>

- Kaltakci-Gurel, D., 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). <https://doi.org/10.1080/02635143.2017.1310094>
- Khoir, M. H., Murtinugraha, R. E., & Musalamah, S. (2020). Pengembangan Media Pembelajaran E-Learning Berbasis Moodle Pada Mata Kuliah Metodologi Penelitian. *Jurnal Pendidikan Teknik Sipil*, 9(1), 54–60.
- Ninawati, M., Burhendi, F. C. A., & Wulandari. (2021). Pengembangan E-Modul Berbasis Software iSpring Suite 9. *Jurnal Educatio*, 7(1). <https://ejournal.unma.ac.id/index.php/educatio/article/view/830/547>
- Rismaningsih, F., & Nurhafsari, A. (2022). Identify Hydrostatic Misconceptions Using Four Tier Diagnostic Tests With the Help of iSpring Suite 9 (Case Study in UNIS Faculty of Engineering Students). *Journal of Science and Science Education (JoSSEd)*, 8(5), 1–4. <https://doi.org/10.29303/jppipa.v1i1.264>
- Rovita, C. A. (2020). *Pengembangan Alat Evaluasi Pembelajaran Matematika Berbasis Two Tier Multiple Choice Menggunakan iSpring Suite 9* [Universitas Muhammadiyah Gresik]. <http://eprints.umg.ac.id/>
- Ruseffendi. (2010). *Dasar-dasar Penelitian Pendidikan & Bidang Non-Eksakta Lainnya*. Tarsito.
- Saftari, M., & Fajriah, N. (2019). Penilaian Ranah Afektif Dalam Bentuk Penilaian Skala Sikap Untuk Menilai Hasil Belajar. *Edutainment : Jurnal Ilmu Pendidikan Dan Kependidikan*, 7(1), 71–81. <https://doi.org/10.35438/e.v7i1.164>
- Sugiyono. (2013). *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)* (1st ed.). Alfabeta.
- Surahman, E., & Surjono, H. D. (2017). Pengembangan adaptive mobile learning pada mata pelajaran biologi SMA sebagai upaya mendukung proses blended learning. *Jurnal Inovasi Teknologi Pendidikan*, 4(1), 26. <https://doi.org/10.21831/jitp.v4i1.9723>
- Tani, S., & Ekawati, E. Y. (2019). Peningkatan Kemandirian Belajar Peserta Didik pada Materi Teori Kinetik Gas Melalui Penerapan Media Pembelajaran Interaktif Berbasis iSpring Suite 8. *Jurnal Materi Dan Pembelajaran Fisika*, 7(2), 13–16. <https://jurnal.uns.ac.id/jmpf/article/view/31454>
- Taslidere, E., Yıldırım, B. (2023). Effect of Conceptual Change–Oriented Instruction on Students' Conceptual Understanding and Attitudes Towards Simple Electricity. *International Journal of Science and Mathematics Education*, 21(1), 1567–1589. <https://doi.org/10.1007/s10763-022-10319-w>
- Wardhono, A., Kalista, A., Kurniawati, D., & Susilo, P. B. (2019). Quiz Training Program through iSpring Suite 8.0 to Junior High School Teachers Tubanitle. *KSILOGIYA : Jurnal Pengabdian Kepada Masyarakat*, 3(1). <https://garuda.kemdikbud.go.id/documents/detail/926494>

Yunita, L., Agung, S., & Noviyanti, Y. (2017). Penerapan Instrumen Penilaian Ranah Afektif Siswa Pada Praktikum Kimia di Sekolah. *Prosiding Seminar Nasional Pendidikan FKIP UNTIRTA*, 1(2), 107–114. <https://jurnal.untirta.ac.id/index.php/psnp/article/view/107-114>