

Pengembangan Digi4t-HB: Four-tier Test Digital untuk Identifikasi Miskonsepsi pada Tekanan Hidrostatik dan Gaya Apung

Development of Digi4T-HB: A Digital Four-Tier Test for Misconception Identification in Hydrostatic Pressure and Buoyancy

Dendy Siti Kamilah¹, Lina Aviyanti^{1*}, Achmad Samsudin¹, Iwan Permana Suwarna²

¹ Universitas Pendidikan Indonesia, Bandung, Indonesia

² UIN Syarif Hidayatullah, Jakarta, Indonesia

Corresponding author: detikamila@upi.edu

ABSTRAK

Latar Belakang: Miskonsepsi siswa pada tekanan hidrostatik dan gaya apung menghambat pemahaman fisika. Meski guru mengidentifikasi miskonsepsi melalui tanya jawab, metode ini kurang efektif. Hasil angket menunjukkan kendala identifikasi miskonsepsi karena waktu dan pemahaman guru terhadap instrumen diagnostik. Penelitian ini bertujuan mengembangkan instrumen *four-tier test* digital (Digi4T-HB) untuk mengidentifikasi miskonsepsi siswa pada materi tekanan hidrostatik dan gaya apung.

Subjek dan Metode: Jenis penelitian ini adalah *Design and Development Research* (DDR). Validitas instrumen diuji dengan menggunakan analisis Rasch Model. Instrumen diujicobakan kepada 51 siswa kelas XI di Tasikmalaya dengan teknik *cluster random sampling*.

Hasil: Hasil validasi konten berdasarkan penilaian ahli yang diolah multifaset menunjukkan semua soal lolos validasi ahli, namun perlu revisi sesuai saran, terutama aspek keterbacaan. validitas konstruk menunjukkan bahwa instrumen memenuhi kriteria unidimensionalitas dengan nilai varians mentah sebesar 34,2% dan nilai Eigen value sebesar 2,2303. Analisis butir soal menghasilkan 12 dari 14 butir soal yang sesuai berdasarkan statistik kecocokan. Tingkat kesulitan pertanyaan bervariasi dengan rentang logit -0,37 hingga +0,50. Uji reliabilitas menunjukkan koefisien alpha Cronbach sebesar 0,82 (kategori sangat baik), reliabilitas orang sebesar 0,72 dan reliabilitas butir soal sebesar 0,71 (keduanya kategori sedang). Hasil dari tahap evaluasi menghasilkan 12 soal DIGI4T-HB yang valid dan reliabel untuk mengidentifikasi miskonsepsi siswa secara digital.

Kesimpulan: Instrumen Digi4T-HB terbukti valid dan reliabel untuk mengidentifikasi miskonsepsi siswa pada topik tekanan hidrostatik dan gaya apung. Penggunaan tes digital meningkatkan efisiensi identifikasi miskonsepsi dan memudahkan guru dalam menganalisis pemahaman siswa. Digi4T-HB dapat digunakan untuk mendukung proses remediasi siswa yang mengalami miskonsepsi.

Kata Kunci: Four-tier Test, Hydrostatic Pressure, Buoyancy, Misconceptions

Korespondensi:

Dendy Siti Kamilah. Universitas Pendidikan Indonesia. Bandung, Indonesia. Email: detikamila@upi.edu Mobile: 08989003374

LATAR BELAKANG

Siswa datang ke kelas dengan pemahaman dan pengalaman terkait materi yang akan dipelajari (Resbiantoro dkk., 2022). Dalam teori konstruktivisme, interpretasi konsep yang dibangun oleh siswa disebut konsepsi (Nasrudin & Azizah, 2020). Konsepsi yang bertentangan dengan teori yang berlaku disebut miskonsepsi (Clement, 1993). Miskonsepsi dalam fisika mengacu pada gagasan atau keyakinan siswa tentang fenomena fisika yang tidak konsisten dengan penjelasan ilmiah yang diterima (Neidorf dkk., 2020). Miskonsepsi berbeda dengan kurangnya pengetahuan (Gurel dkk., 2015). Siswa yang mengalami miskonsepsi tidak menyadari bahwa konsepsi mereka salah, sehingga mereka meyakini bahwa konsepsi tersebut benar (Çelikkani & Kızılcık, 2022). Fluida statis dalam Fisika adalah topik yang memiliki berbagai miskonsepsi. Sub-topik yang memiliki banyak miskonsepsi yang terdeteksi adalah hidrostatik dan gaya apung. Salah satu contoh miskonsepsi adalah penyebab benda terapung dan tenggelam hanya bergantung pada massa atau volume benda (Hunaidah dkk., 2022; Inggit dkk., 2021; Mellu & Langtang, 2023). Miskonsepsi dapat muncul karena berbagai faktor, seperti dari siswa itu sendiri, pendidik, metode pembelajaran, materi pembelajaran, teks referensi, bahasa yang digunakan, dan pengalaman (Resbiantoro dkk., 2022). Miskonsepsi dapat mengganggu proses pembelajaran dan pemahaman konsep sains, terbawa hingga ke jenjang pendidikan yang lebih tinggi, mempengaruhi kemampuan siswa dalam menerapkan konsep sains dalam kehidupan sehari-hari, menghambat kemampuan siswa dalam memecahkan masalah, bahkan menghambat pencapaian keterampilan abad ke-21 (Resbiantoro dkk., 2022; Samsudin dkk., 2024; Vosniadou, 2019). Selain itu, miskonsepsi siswa diyakini meresap, tertanam kuat, dan bertahan

dari waktu ke waktu (Samsudin dkk., 2024). Pendidik perlu mengidentifikasi miskonsepsi sejak dini, karena dapat mengganggu konsep-konsep terkait (Treagust, 2006).

Miskonsepsi dapat diidentifikasi dengan menggunakan tes diagnostik seperti wawancara, pertanyaan terbuka, tes pilihan ganda, dan tes bertingkat (Gurel dkk., 2015; Resbiantoro dkk., 2022; Soeharto dkk., 2019). Wawancara dan pertanyaan terbuka menawarkan keuntungan dalam memberikan wawasan yang mendalam, memungkinkan pertanyaan lanjutan yang mendalam, dan menangkap pemahaman yang bernuansa, tetapi dapat memakan waktu dan membutuhkan interpretasi yang terampil (Resbiantoro dkk., 2022; Soeharto dkk., 2019). Tes pilihan ganda menawarkan efisiensi waktu dan penilaian objektif untuk diagnosis miskonsepsi berskala besar, tetapi tes ini membatasi eksplorasi mendalam terhadap ide-ide siswa dan mungkin tidak mengungkapkan miskonsepsi yang tidak terduga (Resbiantoro dkk., 2022). Tes bertingkat adalah instrumen penilaian yang terdiri dari beberapa tingkat pertanyaan yang dirancang untuk mengevaluasi pemahaman konseptual siswa secara lebih mendalam, dengan berbagai jenis antara lain *two-tier*, *three-tier*, *four-tier*, dan *five-tier tests*, masing-masing secara progresif menambahkan lapisan penilaian seperti penalaran, tingkat kepercayaan diri, dan sumber pengetahuan (Çelikkanlı & Kızılcık, 2022).

Two-tier tests terdiri dari tingkat pertama dengan pertanyaan pilihan ganda dan tingkat kedua yang meminta penalaran, menawarkan keuntungan dalam menilai pemahaman konseptual tetapi memiliki keterbatasan dalam membedakan miskonsepsi dari kurangnya pengetahuan (Soeharto dkk., 2019). *Three-tier tests* terdiri dari pertanyaan konten, pertanyaan penalaran, dan peringkat keyakinan, menawarkan keuntungan yang membedakan miskonsepsi dari kurangnya pengetahuan, tetapi memiliki keterbatasan termasuk meremehkan proporsi kurangnya pengetahuan karena peringkat keyakinan yang terselubung untuk kedua tingkatan (Gurel dkk., 2015). *Four-tier tests* terdiri dari empat tingkat (isi pertanyaan, peringkat keyakinan, alasan jawaban, keyakinan pada alasan), menawarkan akurasi yang lebih baik dalam mengidentifikasi miskonsepsi tetapi membutuhkan lebih banyak waktu untuk mengelola dan menganalisis dibandingkan dengan format yang lebih sederhana (Soeharto dkk., 2019). *Five-tier tests* memberikan informasi yang lebih rinci tentang proses berpikir dan sumber pengetahuan siswa, tetapi bisa jadi terlalu rumit dan memakan waktu bagi peserta tes dan penilai (Çelikkanlı & Kızılcık, 2022). Banyak tingkat pertanyaan dapat menyebabkan peningkatan beban kognitif.

Four-tier tests dipilih karena keseimbangan antara kemampuan diagnostik dan implementasi praktis. *Four-tier test* telah divalidasi dalam berbagai konteks pendidikan, memberikan dasar yang kuat untuk analisis dan perbandingan dengan literatur yang ada. Untuk mengatasi kompleksitas penilaian dan analisis miskonsepsi, tes empat tingkat akan dikembangkan dengan bantuan platform Google Form dan Google Sheet untuk mengotomatiskan pengumpulan jawaban dan data analisis. Sehingga tujuan dari penelitian ini adalah mengembangkan digital *four-tier test* melalui fase analisis, desain, pengembangan, dan evaluasi untuk identifikasi miskonsepsi pada materi tekanan hidrostatik dan gaya apung (Digi4T-HB) yang valid, praktis, dan dapat diterima oleh guru dan siswa.

METODE PENELITIAN

Jenis Penelitian

Pendekatan untuk mengembangkan tes empat tingkat digital menggunakan Design and Development Research (DDR) melibatkan empat Fase yaitu Analisis, Desain, Pengembangan, dan Evaluasi.

1. Analisis

Pada tahap Analisis, penilaian kebutuhan dan identifikasi masalah dilakukan melalui studi literatur dan kuesioner dengan para pendidik fisika. Miskonsepsi yang terjadi pada topik fluida statis dianalisis berdasarkan penelitian terkait dan ditentukan miskonsepsi inti yang perlu dideteksi. Alat diagnostik miskonsepsi yang sudah ada pada topik hidrostatik dan gaya apung dianalisis, seperti *Digital Three-tier Static Fluid Test* (Kamilah & Suwarna, 2016), *Five-tier Fluid Static Test* (5TFST) (Inggit dkk., 2021), dan *Four-tier Static Fluid Test* (Sholahuddin dkk., 2019).

2. Desain

Fase Desain melibatkan pengembangan spesifikasi tes dan cetak biru butir soal yang selaras dengan kurikulum. Set awal dari empat level soal tes dikembangkan berdasarkan soal-soal diagnostik yang telah divalidasi dari penelitian sebelumnya. Antarmuka digital dan analisis miskonsepsi dirancang.

3. Pengembangan

Pada tahap Pengembangan, prototipe tes empat tingkat digital divalidasi oleh lima orang ahli untuk validitas konten. Setelah direvisi, tes tersebut didigitalkan dengan menggunakan Google Form. Uji coba skala kecil dengan siswa menilai desain antarmuka, kegunaan, dan ketertarikan terhadap format.

4. Evaluasi

Pada tahap Evaluasi, tes empat tingkat digital yang telah selesai diberikan kepada sampel siswa yang lebih besar. Sifat psikometrik, validitas konstruk, dan reliabilitas dianalisis menggunakan Rasch Model. Selain itu, dasbor analisis miskonsepsi otomatis yang dikembangkan dengan Google Sheet diperkenalkan kepada para guru fisika. Mereka menjelajahi dasbor, menginterpretasikan hasil, dan memberikan umpan balik melalui kuesioner. Tahap ini

menilai kegunaan dan interpretabilitas dari sudut pandang guru dan mengumpulkan saran perbaikan Digi4T-HB selanjutnya disempurnakan berdasarkan hasil tersebut.

Sampel dan Populasi

Populasi dalam penelitian ini adalah seluruh siswa kelas 11 pada salah satu sekolah di Tasikmalaya. Sampel penelitian untuk uji coba skala kecil adalah 7 siswa kelas XI secara *purposive sampling* dan sample untuk uji coba skala besar adalah 51 siswa dengan *cluster random sampling*. Instrumen yang digunakan kuesioner dan *four-tier test*. Kuesioner digunakan untuk analisis kebutuhan guru, respon siswa terhadap Digi4T-HB, dan tanggapan guru terhadap tampilan analisis miskonsepsi pada Google Sheet. Teknik analisis data yang digunakan dalam penelitian ini meliputi analisis Rasch Model untuk menguji validitas konstruk, reliabilitas, dan karakteristik butir soal dari instrumen Digi4T-HB.

HASIL PENELITIAN

Hasil penelitian disajikan berdasarkan temuan pada setiap fase.

Fase Analisis

Hasil identifikasi masalah melalui kuesioner kepada 17 guru dengan menggunakan Google Form disajikan pada Tabel 1. Kuesioner guru menunjukkan bahwa guru tidak melakukan tes khusus untuk mengidentifikasi ada tidaknya miskonsepsi pada siswa. Keterbatasan waktu dan kurangnya pengetahuan dalam membuat tes diagnostik miskonsepsi menjadi faktor utama guru mengapa tidak melakukan identifikasi miskonsepsi pada seluruh siswa.

Tabel 1. Hasil kuesioner analisis kebutuhan guru

Pertanyaan kuesioner	Tanggapan terbanyak
Apakah Anda pernah memeriksa atau memastikan tidak ada miskonsepsi di antara para siswa?	Already (12 respondents)
Bagaimana Anda memastikan bahwa peserta didik tidak memiliki miskonsepsi?	1. Tanya jawab (9 responden) 2. Hasil penugasan (6 responden)
Pada konsep apa saja siswa sering mengalami miskonsepsi? (sebutkan 3)	1. Fluida statis (7 responden) 2. Dinamika (6 responden)
Apakah guru pernah menggunakan instrumen khusus untuk mengidentifikasi miskonsepsi?	Tidak pernah (17 responden)
* jika tidak	1. Tidak ada waktu (11 respondents)
Apa alasan guru tidak menggunakan instrumen diagnostik miskonsepsi?	2. Tidak mengerti (6 respondents)

Fase Desain

Pada tahap desain, sembilan miskonsepsi utama yang ditemukan didistribusikan ke dalam 14 butir soal (lihat Tabel 2). Digi4T-HB dikembangkan dari beberapa sumber soal yang telah divalidasi termasuk Kamilah & Suwarna (2016), Sholahuddin (2019), dan Inggit dkk. (2021).

Tabel 2. Distribusi Soal

Kode	Miskonsepsi	No. Soal
Tekanan Hidrostatik		
M1	Tekanan hidrostatik dipengaruhi oleh luas penampang wadah (Djudin, 2021; Inggit dkk., 2021; Kamilah, 2016; Karaman, 2011; Maknun & Marwiah, 2022; Sholahuddin dkk., 2019)	1 and 2
M2	Kedalaman yang mempengaruhi tekanan hidrostatik bukanlah kedalaman dari permukaan fluida (Inggit dkk., 2021; Sholahuddin, 2019).	3 and 4
M3	Tekanan hidrostatik tidak dipengaruhi oleh massa jenis fluida fluida (Sholahuddin dkk., 2019).	5
M4	Fluida tidak memberikan tekanan ke segala arah dan sama besarnya (Kamilah, 2016).	6
Gaya Apung		
M5	Besarnya gaya apung dipengaruhi oleh massa benda, densitas benda, atau volume benda (Djudin, 2021; Inggit dkk., 2021; Kamilah, 2016).	7 and 8
M6	Besar gaya apung dipengaruhi oleh besar volume benda, luas permukaan yang bersentuhan dengan fluida, atau massa jenis benda (Djudin, 2021; Inggit dkk., 2021; Kamilah, 2016; Sholahuddin, 2019).	9
M7	Tenggelam/terapungnya suatu benda dalam fluida ditentukan oleh volume, massa, atau senyawanya (Kamilah, 2016; Maknun & Marwiah, 2022).	10 and 11
M8	Tenggelam/terapungnya suatu benda dalam fluida ditentukan oleh volume fluida, massa fluida, atau luas penampang zat cair (Kamilah, 2016) (Maknun & Marwiah, 2022).	12 and 13
M9	Berat benda yang terukur di dalam air dipengaruhi oleh volume air, massa air, atau kekentalan air (Kamilah, 2016).	14

Fase Pengembangan

1. Validasi konten oleh ahli

Pada fase pengembangan, kisi-kisi soal dibuat dan divalidasi oleh para ahli untuk menilai validitas konten. Lima orang ahli (dua dosen dan tiga guru) menilai aspek kurikulum, keterbacaan, dan konsep untuk masing-masing item soal. Indikator untuk setiap aspek disajikan pada Tabel 3.

Tabel 3. Aspek dan indikator validitas ahli

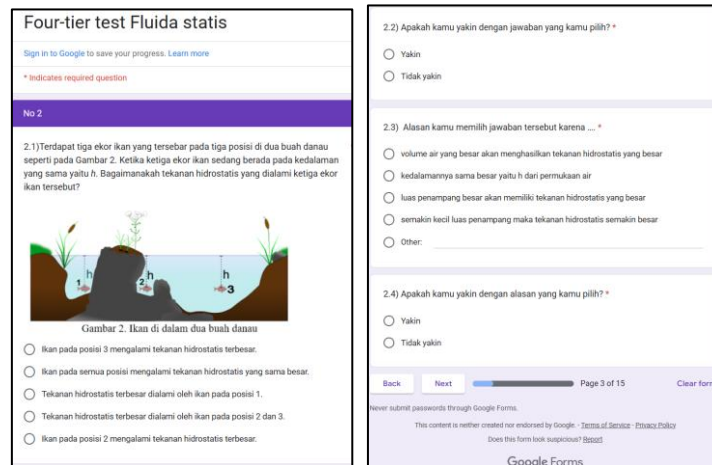
Indikator	Kode
Kurikulum	
Hanya menguji konsep yang sesuai dengan jenjang pendidikan yang ditargetkan	JenjangPendidikan
Soal disusun berdasarkan materi yang benar-benar diajarkan pada kurikulum yang berlaku	Kurikulum
Keterbacaan	
Kalimat tidak ambigu dan sesuai dengan kaidah bahasa yang baik dan benar	KaidahBahasa
Tidak menggunakan kalimat yang terlalu panjang atau kompleks.	KalimatEfektif
Pernyataan dalam soal harus bersifat netral dan tidak mengarahkan jawaban.	TidakMengarahJawaban
Konsep	
Kesesuaian antara teori ilmiah dan contoh kasus yang diberikan	KesesuaianTeori
Hanya ada satu kunci jawaban	Hanya1Jawaban
Instrumen hanya menguji konsep yang berkaitan dengan miskonsepsi	SesuaiMiskonsepsi

Hasil validasi dianalisis dengan pemodelan Rasch menggunakan multivariate dengan perangkat lunak Minifac. Hasil uji validitas oleh para ahli ditunjukkan pada Gambar 1. Gambar 1 menunjukkan distribusi pertanyaan di sebelah kiri, distribusi indikator validasi di tengah, dan distribusi ahli di sebelah kanan.



Gambar 1. Wright Map hasil analisis multifacet validasi ahli

Instrumen yang telah valid dan direvisi kemudian dikonversi ke dalam bentuk digital menggunakan Google Form (lihat Gambar 2).



Four-tier test Fluida statis

Sign in to Google to save your progress. [Learn more](#)

* Indicates required question

No 2

2.1) Terdapat tiga ekor ikan yang tersebar pada tiga posisi di dua buah danau seperti pada Gambar 2. Ketika ketiga ekor ikan sedang berada pada kedalaman yang sama yaitu h . Bagaimanakah tekanan hidrostatik yang dialami ketiga ekor ikan tersebut?



Gambar 2. Ikan di dalam dua buah danau

☐ Ikan pada posisi 3 mengalami tekanan hidrostatik terbesar.

☐ Ikan pada semua posisi mengalami tekanan hidrostatik yang sama besar.

☐ Tekanan hidrostatik terbesar dialami oleh ikan pada posisi 1.

☐ Tekanan hidrostatik terbesar dialami oleh ikan pada posisi 2 dan 3.

☐ Ikan pada posisi 2 mengalami tekanan hidrostatik terbesar.

2.2) Apakah kamu yakin dengan jawaban yang kamu pilih? *

☐ Yakin

☐ Tidak yakin

2.3) Alasan kamu memilih jawaban tersebut karena ... *

☐ volume air yang besar akan menghasilkan tekanan hidrostatik yang besar

☐ kedalamannya sama besar yaitu h dari permukaan air

☐ luas penampang besar akan memiliki tekanan hidrostatik yang besar

☐ semakin kecil luas penampang maka tekanan hidrostatik semakin besar

☐ Other: _____

2.4) Apakah kamu yakin dengan alasan yang kamu pilih? *

☐ Yakin

☐ Tidak yakin

Back Next Page 3 of 15 Clear form

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. [Terms of Service](#) [Privacy Policy](#)

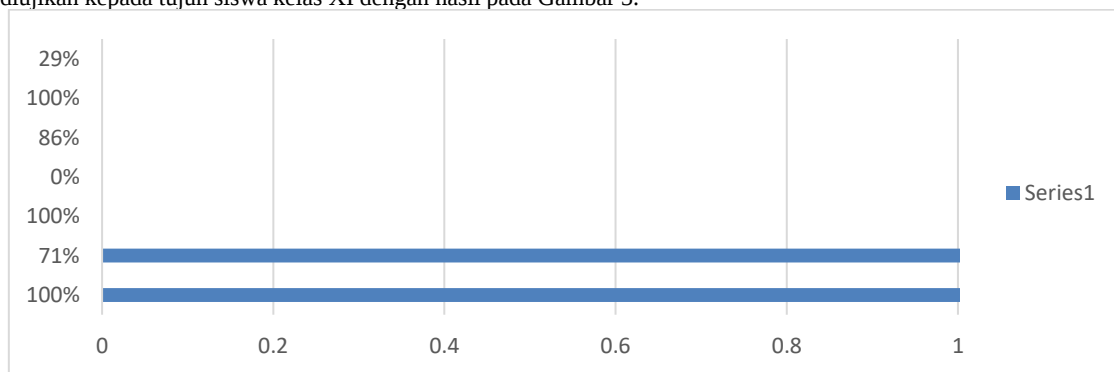
Does this form look suspicious? [Report](#)

Google Forms

Gambar 2. Tampilan Contoh Soal Digi4T-HB

2. Uji coba skala kecil

Uji coba skala kecil dilakukan untuk mengetahui perlunya perbaikan dalam fase pengembangan. Digi4T-HB diujikan kepada tujuh siswa kelas XI dengan hasil pada Gambar 3.



Gambar 3. Hasil kuesioner respon siswa

Fase Evaluasi

Uji coba skala besar dilakukan untuk tahap evaluasi. Tahap ini memastikan validitas konstruk, reliabilitas, dan analisis butir soal (*fit statistic* dan tingkat kesukaran butir soal).

1. Validitas Konstruk

Uji validitas konstruk dilakukan dengan analisis unidimensionalitas menggunakan Rasch Model. Pada Gambar 5 diketahui nilai *raw variance by measures* adalah 34,2%, sedangkan pada *raw unexplained variance 1st contrast* memiliki *eigenvalues* 2,2303 dan *observed* 10,15%.

INPUT: 51 Person 14 Item REPORTED: 51 Person 14 Item 5 CATS MINISTEP 5.9.0.0

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

	Eigenvalue	Observed	Expected
Total raw variance in observations =	21.2739	100.0%	100.0%
Raw variance explained by measures =	7.2739	34.2%	34.5%
Raw variance explained by persons =	5.8593	27.5%	27.8%
Raw Variance explained by items =	1.4146	6.6%	6.7%
Raw unexplained variance (total) =	14.0000	65.8%	65.5%
Unexplned variance in 1st contrast =	2.2303	10.5%	15.9%
Unexplned variance in 2nd contrast =	2.0536	9.6%	14.5%
Unexplned variance in 3rd contrast =	1.7522	8.2%	12.5%
Unexplned variance in 4th contrast =	1.4926	7.0%	10.7%
Unexplned variance in 5th contrast =	1.2711	6.0%	9.1%

Essential Unidimensionality (Rasch/Common variance) = 49.2%

Gambar 5. Standardized Residual

2. Uji Reliabilitas

Analisis reliabilitas dilakukan dengan bantuan aplikasi Ministep, hasil terlihat pada Tabel 3. Koefisien *alpha Cronbach* bernilai sebesar 0,82. Sedangkan person reliability bernilai 0,72 dan item reliability memiliki nilai 0,71.

Tabel 3. Hasil uji reliabilitas

Reliabilitas	Nilai	Kategori
<i>Cronbach Alpha</i> (KR-20)	0,82	Sangat baik
<i>Person Reliability</i>	0,72	Cukup
<i>Item Reliability</i>	0,71	Cukup

3. Analisis butir soal

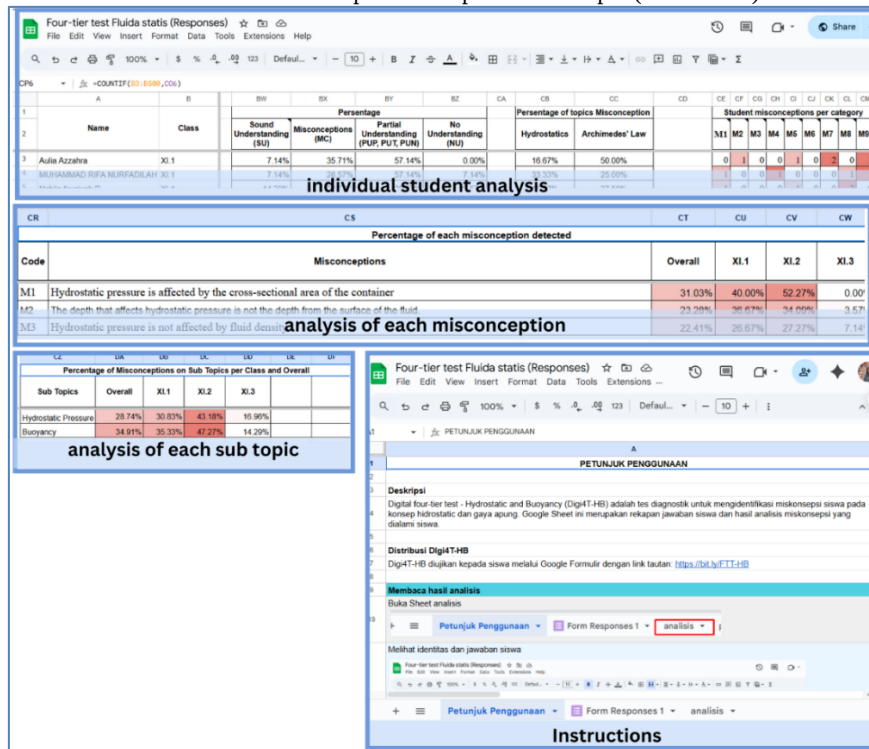
Fit statistics yang digunakan untuk menentukan validitas konstruk adalah infit mean-square (MNSQ), outfit MNSQ, dan point measure correlation (Pt. Measure Corr). Tingkat kesulitan butir soal (item difficulty) dianalisis menggunakan model Rasch berdasarkan nilai Logarithm odd unit (Logit). Hasil analisis butir soal ditunjukkan pada Tabel 4 yang terdiri dari nilai Infit dan Outfit MNSQ, Pt. Measure Corr dan Logit serta interpretasinya..

Tabel 4. Hasil analisis butir soal

Item Soal	Fit Statistics			Tingkat Kesulitan		
	Infit MNSQ	Outfit MNSQ	Pt-Mea Corr	Interpretasi	Logit	Interpretasi
1	0,88	0,85	0,62	Sangat sesuai	-0,23	Sedang
2	0,83	0,75	0,81	Sangat sesuai	-0,32	Mudah
3	0,93	0,87	0,62	Sangat sesuai	-0,37	Mudah
4	0,81	0,72	0,60	Sangat sesuai	0,00	Sedang
5	1,21	1,29	0,32	Sesuai	-0,21	Sedang
6	1,67	1,60	0,33	Tidak sesuai	0,06	Sedang
7	1,13	0,96	0,38	Sesuai	0,27	Sulit
8	1,51	1,37	0,23	Kurang sesuai	0,50	Sulit
9	0,76	0,87	0,42	Sangat sesuai	0,02	Sedang
10	0,97	0,99	0,54	Sangat sesuai	-0,15	Sedang
11	0,87	0,72	0,75	Sangat sesuai	0,27	Sulit
12	0,92	0,95	0,53	Sangat sesuai	0,02	Sedang
13	0,81	0,83	0,57	Sangat sesuai	0,03	Sedang
14	1,02	0,87	0,67	Sangat sesuai	0,12	Sedang

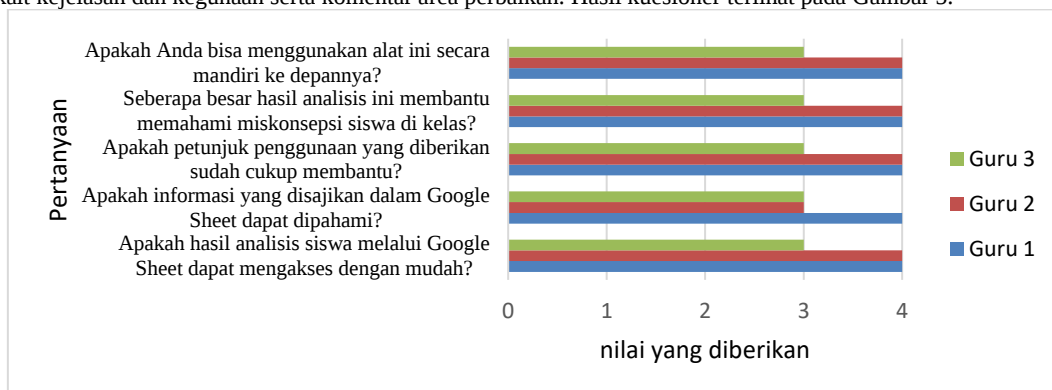
4. Penilaian guru terhadap tampilan analisis

Setelah siswa menyelesaikan tes diagnostik melalui Google Formulir, jawaban mereka secara otomatis diproses dan divisualisasikan dalam dasbor Google Sheet yang terstruktur. Dasbor ini mencakup analisis siswa secara individu, analisis keseluruhan dan analisis kelas untuk setiap miskonsepsi dan sub-topik (lihat Gbr. 4).



Gambar 4. Tampilan Analisis Google Sheet

Untuk mengevaluasi kegunaan dan kejelasan alat analisis ini, kuesioner umpan balik didistribusikan kepada guru. Para guru diminta menjelajahi antarmuka Google Sheet, memahami analisis, dan memberikan penilaian 1-4 terkait kejelasan dan kegunaan serta komentar area perbaikan. Hasil kuesioner terlihat pada Gambar 5.



Gambar 5. Hasil kuesioner guru terhadap Google Sheet

PEMBAHASAN

Hasil penelitian akan dibahas berdasarkan fase-fase DDR. Setiap fase diuraikan secara sistematis, menjelaskan temuan-temuan penting, tantangan yang dihadapi, serta solusi yang diterapkan.

Fase Analisis

Berdasarkan hasil analisis kebutuhan guru diperlukan alat tes diagnostik miskonsepsi massal dengan kemudahan dalam menganalisis hasil diagnostik, sehingga guru dapat mengidentifikasi miskonsepsi seluruh siswa. Masalah ini dapat diatasi dengan menggunakan tes digital sehingga analisis dapat dilakukan secara otomatis oleh sistem. Tes digital dalam pendidikan mengacu pada penggunaan alat dan platform digital untuk mengelola, memberi skor, dan mengevaluasi penilaian pembelajaran siswa (Khan dkk., 2022). Topik fisika yang memiliki urgensi untuk dikembangkan instrumen diagnostik miskonsepsinya adalah fluida statis khususnya subtopik tekanan hidrostatik dan gaya apung (Firdausi dkk., 2019; Handayani dkk., 2022; Hidayat dkk., 2022; Ozkan & Selcuk, 2015)

Fase Desain

Fase desain merupakan tahap kritis dalam proses pengembangan instrumen pada penelitian ini, di mana miskonsepsi dan soal-soal terkait diidentifikasi dan dirumuskan. Miskonsepsi yang tercantum dalam Tabel 2 merupakan hasil analisis yang didasarkan pada miskonsepsi yang sering terjadi pada subtopik tekanan hidrostatik dan gaya apung, sebagaimana diidentifikasi dalam penelitian-penelitian sebelumnya. Google Formulir dipilih sebagai platform pendistribusian soal, karena Google Formulir menyediakan platform yang efisien untuk mengumpulkan respons siswa dan melakukan penilaian dalam format digital (Park dkk., 2021). Selanjutnya Google Formulir secara otomatis mengumpulkan respons dalam Google Sheet yang ditautkan, sehingga memungkinkan transfer dan analisis data tanpa hambatan (Arulogun dkk., 2020).

Fase Pengembangan

Pada fase pengembangan, instrumen Digi4T-HB melalui dua tahap utama yaitu validitas konten oleh ahli dan uji coba skala kecil. Validitas konten adalah tingkat sejauh mana suatu instrumen pengukuran mencakup dan mewakili seluruh aspek atau domain yang hendak diukur secara komprehensif dan relevan (Çelikkanlı, 2022). Ahli yang terdiri dari dosen dan guru fisika menilai setiap item untuk memastikan representasi yang tepat dari konten yang diharapkan, serta kejelasan dan keakuratan dalam penyajian pertanyaan dan pilihan jawaban (Kaltakçı, 2012). Data dalam kuesioner validasi dikonversi dengan memberikan skor 3 untuk "valid tanpa revisi", skor 2 untuk "valid dengan revisi", dan skor 1 untuk "tidak valid". Item soal nomor 10, 12, 13, dan 9 dinilai valid tanpa revisi pada semua indikator dengan nilai logit 3. Indikator "Kaidah Bahasa" merupakan indikator yang paling sulit dipenuhi, sehingga item soal yang berada dalam kotak berwarna merah pada Gambar 1 direvisi terkait kaidah bahasa. Indikator-indikator di bagian bawah diagram dinilai valid tanpa revisi oleh semua validator, mencakup aspek kurikulum dan konten (lihat Tabel 3). Hanya aspek keterbacaan yang perlu direvisi oleh semua validator. Di sebelah kanan diagram, "validator4" memiliki penilaian yang ketat dengan logit -2, namun selaras dengan pertanyaan nomor empat yang memiliki nilai logit mendekati -2. Hal ini mengindikasikan bahwa semua pertanyaan lulus validasi ahli, namun perlu revisi seperti yang disarankan.

Hasil uji coba skala kecil menunjukkan Google formulir mudah digunakan. Kuesioner didistribusikan mengenai aksesibilitas, kejelasan visual, dan pengalaman pengguna. Hasil kuesioner dapat dilihat pada Gambar 3. Mayoritas siswa menyatakan tes ini mudah digunakan, dengan 100% responden setuju bahwa tes mudah diakses di perangkat mereka. Hal ini didukung karena siswa terbiasa menggunakan Google Formulir untuk assesmen. Zitha dkk. (2023) menemukan bahwa siswa akan mengalami kesulitan menavigasi platform daring berbeda jika platform tersebut baru dan rumit. Semua siswa menyatakan pertanyaan dan gambar ditampilkan dengan jelas, mengindikasikan desain antarmuka sudah sesuai. Mengenai minat, 85,7% siswa lebih suka menggunakan Google Formulir daripada format tradisional. Meskipun beberapa siswa (28,6%) menyarankan perbaikan tata letak visual, tidak ada yang melaporkan kesulitan mengoperasikan Google Formulir. Kendala hanya terkait konektivitas internet yang dialami satu dari tujuh siswa. Para siswa juga memberikan umpan balik terbuka. Mayoritas merasa tes ini mudah dan dirancang dengan baik, namun beberapa menyarankan penyederhanaan kata dalam pertanyaan, terutama istilah ilmiah seperti "bejana". Mereka juga merekomendasikan peningkatan kejelasan diagram dan bahasa dalam pilihan jawaban untuk meningkatkan kenyamanan tes.

Fase Evaluasi

Pada fase evaluasi, instrumen Digi4T-HB diuji lebih lanjut pada uji coba skala besar untuk memastikan validitas, reliabilitas, dan analisis butir soal. Selain itu antarmuka analisis digital pada Google Sheet juga diperkenalkan kepada guru untuk evaluasi praktis. Langkah ini bertujuan untuk menilai apakah guru dapat secara efektif menafsirkan dan mengambil manfaat dari analisis miskonsepsi yang disediakan.

1. Validitas konstruk

Validitas konstruk penting untuk memastikan bahwa instrumen penelitian benar-benar mengukur apa yang dimaksudkan untuk diukur (Heale & Twycross, 2015). Validitas konstruk dari *four-tier test* dapat dianalisis dengan *Rasch model* (Bozdağ & Türkoğuz, 2021). Analisis unidimensionalitas bertujuan untuk memastikan bahwa semua item dalam tes mengukur sifat laten yang sama atau mengukur apa yang seharusnya diukur (Ishak dkk., 2018). Sebuah instrumen dikatakan valid jika nilai raw variance data minimal 20% dan raw unexplained variance kontras pertama memiliki nilai eigenvalues di bawah 3,00 dan nilai observasi kurang dari 15% (Kunaedi dkk., 2024). Tiga persyaratan

unidimensionalitas terpenuhi, yang mengindikasikan instrumen fluida statis digital uji empat tingkat ini dapat mengukur apa yang seharusnya diukur.

2. Uji reliabilitas

Reliabilitas adalah ukuran yang menunjukkan sejauh mana item-item dalam tes konsisten dalam mengukur kemampuan individu (Sukarelawan dkk., 2024). Uji reliabilitas *four-tier test* dilakukan dengan menghitung koefisien alpha Cronbach untuk memeriksa konsistensi internal (Çelikkanlı, 2022; Kaltakçi, 2012). Koefisien Cronbach's alpha sebesar 0,82 termasuk dalam kategori "Sangat Baik", yang mengindikasikan konsistensi internal yang tinggi dari instrumen (Saidi & Siew, 2019). Hal ini menunjukkan bahwa item-item dalam skala ini terkait erat dan mengukur konstruk dasar yang sama. Nilai *person reliability* sebesar 0.72 dan nilai *item reliability* sebesar 0.71 keduanya termasuk dalam kategori "Cukup" (Saidi & Siew, 2019). *Person reliability* menunjukkan konsistensi urutan orang yang dapat diharapkan jika sampel orang ini diberikan satu set item lain yang mengukur konstruk yang sama. *Item reliability* mengacu pada replikasi penempatan item di sepanjang skala jika item yang sama diberikan kepada sampel lain dengan tingkat kemampuan yang sebanding. Menariknya, meskipun Cronbach's alpha menunjukkan konsistensi internal yang sangat baik, nilai reliabilitas orang dan item agak lebih rendah, meskipun masih dapat diterima. Perbedaan ini dapat menunjukkan bahwa meskipun item-item tersebut konsisten secara internal, mungkin ada beberapa variabilitas dalam cara individu merespons item-item tersebut atau bagaimana item-item tersebut mendiskriminasikan antara berbagai tingkat sifat yang diukur. Penting untuk diperhatikan bahwa ukuran reliabilitas dapat dipengaruhi oleh faktor-faktor seperti ukuran sampel, tingkat kesulitan butir soal, dan rentang kemampuan dalam sampel (Devitt dkk., 1998).

3. Analisis butir soal

Dari 14 butir soal yang dianalisis terdapat dua butir soal yang perlu dipertimbangkan untuk digunakan dalam mengukur apa yang seharusnya diukur. Soal nomor enam masuk kategori "tidak sesuai" karena nilai Infit MNSQ, Outfit MNSQ, dan Pt. Measure Corr tidak sesuai kriteria. Oleh karena itu, soal nomor enam sebaiknya tidak digunakan. Soal nomor delapan juga perlu dipertimbangkan karena nilai Infit MNSQ dan Pt. Measure Corr tidak sesuai kriteria sehingga masuk kategori "kurang sesuai". Butir soal kategori "kurang sesuai" masih dapat digunakan dengan revisi. Namun karena indikator miskonsepsi pada soal nomor delapan sudah terwakili oleh butir soal lain, soal ini dapat dihilangkan. Nilai Pt. Measure Corr pada semua butir soal bernilai positif menunjukkan bahwa item tersebut selaras dengan variabel laten yang diukur (Zynuddin & Sumintono, 2024). Rentang logit yang diperoleh adalah -0,37 hingga +0,50 dengan SD 0,24. Standar deviasi yang lebih kecil menunjukkan stabilitas instrumen, didukung dengan tidak ada item yang memiliki nilai logit ekstrem. Terdapat tiga butir soal yang masuk kategori sulit yaitu nomor 7, 8, dan 11. Butir soal yang masuk kategori mudah adalah nomor 2 dan 3 dengan logit 0,32 dan 0,37. Sedangkan butir soal lainnya masuk kategori sedang.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa instrumen Digi4T-HB (Digital Four-tier Test for Hydrostatic Pressure and Buoyancy) yang dikembangkan terbukti valid dan reliabel untuk mengidentifikasi miskonsepsi siswa pada topik tekanan hidrostatik dan gaya apung. Penggunaan format tes digital meningkatkan efisiensi proses identifikasi miskonsepsi dan memudahkan guru dalam menganalisis pemahaman konsep siswa. Instrumen Digi4T-HB yang terdiri dari 12 soal valid ini dapat digunakan untuk mendukung proses remediasi siswa yang mengalami miskonsepsi pada materi tersebut.

DAFTAR PUSTAKA

- Arulogun, O. T., Akande, O. N., Akindele, A. T., & Badmus, T. A. (2020). Survey dataset on open and distance learning students' intention to use social media and emerging technologies for online facilitation. *Data in Brief*, 31, 105929. <https://doi.org/10.1016/J.DIB.2020.105929>
- Bozdağ, C. H., & Türkoğuz, S. (2021). A Rasch Model Analysis of Primary School Students' Conceptual Understanding Levels of The Concept of Light. *International Online Journal of Primary Education (IOJPE)*, 10(1), 160–179. www.iojpe.orgORCID:<https://orcid.org/0000-0001-6735-7096>ORCID:<https://orcid.org/0000-0002-7850-2305>
- Çelikkanlı, N. Ö. (2022). A Review of Studies About Four-Tier Diagnostic Tests in Physics Education. *Journal of Turkish Science Education*, 19(4), 1291–1311. <https://doi.org/10.36681/tused.2022.175>
- Çelikkanlı, N. Ö., & Kızılcık, H. Ş. (2022). A Review of Studies About Four-Tier Diagnostic Tests in Physics Education. *Journal of Turkish Science Education*, 19(4), 1291–1311. <https://doi.org/10.36681/tused.2022.175>
- Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching*, 30(10), 1241–1257. <https://doi.org/10.1002/tea.3660301007>

- Devitt, J. H., Kurrek, M. M., Cohen, M. M., Fish, K., Fish, P., Noel, A. G., & Szalai, J.-P. (1998). Testing Internal Consistency and Construct Validity During Evaluation of Performance in a Patient Simulator. *Anesthesia & Analgesia*, 86(6), 1160–1164. <https://doi.org/10.1213/00000539-199806000-00004>
- Djudin, T. (2021). Promoting Students' Conceptual Change by Integrating The 3-2-1 Reading Technique with Refutation Text in The Physics Learning of Buoyancy. *Journal of Turkish Science Education*, 18(2), 290–303. <https://doi.org/10.36681/tused.2021.66>
- Firdausi, F., Yuliati, L., & Parno, P. (2019). Students' conceptual errors on hydrostatics pressure and bouyancy theory. *Jurnal Pendidikan Sains Universitas ...*. <https://www.neliti.com/publications/477971/students-conceptual-errors-on-hydrostatics-pressure-and-bouyancy-theory>
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(5), 989–1008. <https://doi.org/10.12973/eurasia.2015.1369a>
- Handayani, D. Y., Suhandi, A., & ... (2022). The development of computer supported conceptual change text on buoyant force. *Jurnal Multidisiplin* <https://journal.formosapublisher.org/index.php/mudima/article/view/1542>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. Dalam *Evidence-Based Nursing* (Vol. 18, Nomor 3, hlm. 66–67). BMJ Publishing Group. <https://doi.org/10.1136/eb-2015-102129>
- Hidayat, M., Kurniawan, D. A., Sandra, R. O., & ... (2022). Misconception Tool: Web-Based Assessment of Buoyancy Materials. *Journal of ...*. <https://ejournal.undiksha.ac.id/index.php/JET/article/view/43749>
- Hunaidah, H., Erniwati, E., & Mahdiannur, M. A. (2022). Tes Diagnostik Four-tier untuk Menilai Miskonsepsi Siswa tentang Fluida: Sebuah Kisah dari Pengembangan hingga Pengukuran dari Tiga Lokasi Lingkungan. *Jurnal Penelitian Pendidikan IPA*, 8(3), 1586–1592. <https://doi.org/10.29303/jppipa.v8i3.1784>
- Inggit, S. M., Liliawati, W., & Suryana, I. (2021). Identifikasi Miskonsepsi dan Penyebabnya Menggunakan Instrumen Five-Tier Fluid Static Test (5TFST) pada Peserta Didik Kelas XI Sekolah Menengah Atas. *Journal of Teaching and Learning Physics*, 6(1), 49–68. <https://doi.org/10.15575/jotalp.v6i1.11016>
- Ishak, A. H., Osman, M. R., Mahaiyadin, M. H., Tumiran, M. A., & Anas, N. (2018). Examining Unidimensionality of Psychometric Properties Via Rasch Model. *International Journal of Civil Engineering and Technology (IJCIET)*, 9(9), 1462–1467. <http://www.iaeme.com/IJCIET/index.asp1462http://www.iaeme.com/ijciet/issues.asp?JType=IJCIET&VType=9&IType=9http://www.iaeme.com/IJCIET/index.asp1463http://www.iaeme.com/IJCIET/issues.asp?JType=IJCIET&VType=9&IType=9>
- Kaltakçi, D. (2012). *Development And Application Of A Four-Tier Test To Assess Pre-Service Physics Teachers' Misconceptions About Geometrical Optics A Thesis Submitted To The Graduate School Of Natural And Applied Sciences Of Middle East Technical University*. Middle East Technical University.
- Kamilah, D. S. (2016). Pengembangan three-tier test digital untuk mengidentifikasi miskonsepsi pada konsep fluida statis [UIN Syarif Hidayatullah Jakarta]. Dalam *academia.edu*. https://www.academia.edu/download/70802922/DENDY_20SITI_20KAMILAH-FITK.pdf
- Kamilah, D. S., & Suwarna, I. P. (2016). Pengembangan Three-Tier Test Digital untuk Mengidentifikasi Miskonsepsi Pada Konsep Fluida Statis. *EDUSAINS*, 8(2), 212–220. <https://doi.org/10.15408/es.v8i2.5192>
- Karaman, I. (2011). Effect of instruction based on conceptual change text on students' understanding of fluid pressure concept. Dalam *Int. J. Innovation and Learning* (Vol. 9, Nomor 1).
- Khan, H., Kushwah, K. K., Khare, K., Mahobia, S. K., & Thakur, J. S. (2022). *Methods of Evaluation and Assessment in Digital Learning* (hlm. 83–110). <https://doi.org/10.4018/978-1-6684-5400-8.ch004>
- Kunaedi, J., Samsudin, A., Hasanah, L., Hidiiana Aminudin, A., Herliyana Dewi, F., Alfionita Umar, F., Rona Wahyu Astuti, I., & Nurul Mufida, S. (2024). Developing Diagnostic Test on Geometrical Optics (DT-GO) Concept. *International Conference On Mathematics And Science Education*, 712. <https://doi.org/10.18502/kss.v9i8.15637>
- Maknun, J., & Marwiah, M. (2022). Remediation of Misconceptions Vocational High School Students on the Concept of Static Fluids using the Conceptual Change Model. *Journal of Technical Education and Training*, 14(2 Special Issue), 49–56. <https://doi.org/10.30880/jtet.2022.14.02.005>
- Mellu, R. N. K., & Langtang, D. (2023). Profil Miskonsepsi Peserta Didik pada Materi Kinematika Gerak dan Fluida Statis. *Jurnal Pendidikan Fisika*, 11(2), 170. <https://doi.org/10.24127/jpf.v11i2.7345>
- Nasrudin, H., & Azizah, U. (2020). Overcoming Misconception In Energetic Topics Through Implementation Of Metacognitive Skills-Based Instructional Materials: A Case Study In Student Of Chemistry Department, Universitas Negeri Surabaya. *JPII*, 9(1), 125–134. <https://doi.org/10.15294/jpii.v9i1>
- Neidorf, T., Arora, A., Erberber, E., Tsokodayi, Y., & Mai, T. (2020). *Student Misconceptions and Errors in Physics and Mathematics*. 9. <https://doi.org/10.1007/978-3-030-30188-0>
- Ozkan, G., & Selcuk, G. S. (2015). Effect of Technology Enhanced Conceptual Change Texts on Students' Understanding of Buoyant Force. *Universal Journal of Educational Research*, 3(12), 981–988. <https://doi.org/10.13189/ujer.2015.031205>

- Park, J. J., Edward, H. J. K., & Chung, C. W. (2021). Innovative digital tools for new trends in teaching and assessment methods in medical and dental education. *Journal of Educational Evaluation for Health Professions*, 18. <https://doi.org/10.3352/JEEHP.2021.18.13>
- Resbiantoro, G., Setiani, R., & Dwikoranto. (2022). A Review of Misconception in Physics: The Diagnosis, Causes, and Remediation. *Journal of Turkish Science Education*, 19(2), 403–427. <https://doi.org/10.36681/tused.2022.128>
- Saidi, S. S., & Siew, N. M. (2019). Reliability and Validity Analysis of Statistical Reasoning Test Survey Instrument using the Rasch Measurement Model. *International Electronic Journal of Mathematics Education*, 14(3), 535–546. <https://doi.org/10.29333/IEJME/5755>
- Samsudin, A., Azizah, N., Fratiwi, N. J., Suhandi, A., Irwandani, I., Nurtanto, M., Yusup, M., Supriyatman, S., Masrifah, M., Aminudin, A. H., & Costu, B. (2024). Development of DIGaKiT: identifying students' alternative conceptions by Rasch analysis model. *Journal of Education and Learning*, 18(1), 128–139. <https://doi.org/10.11591/edulearn.v18i1.20970>
- Sholahuddin, S. (2019). *Penerapan Conceptual Change Texts (Cct) Berbantuan Komputer Untuk Menurunkan Miskonsepsi Siswa Pada Materi Fluida Statis*. repository.upi.edu. <http://repository.upi.edu/id/eprint/48321>
- Sholahuddin, S., Rusnayani, H., & Suyana, I. (2019). Profil Miskonsepsi Siswa Berdasarkan Hasil Tes Diagnostik Four Tier Test pada Materi Fluida Statis. *WaPFI (Wahana Pendidikan Fisika)*, 4(2), 194–199. <https://doi.org/10.17509/wapfi.v4i2.20201>
- Soeharto, Csapó, B., Sarimanah, E., Dewi, F. I., & Sabri, T. (2019). A review of students' common misconceptions in science and their diagnostic assessment tools. *Jurnal Pendidikan IPA Indonesia*, 8(2), 247–266. <https://doi.org/10.15294/jpii.v8i2.18649>
- Sukarelawan, M. I., Indratno, T. K., Oktova, R., & ... (2024). Quality of Four-Tier Diagnostic Test on Wave and Vibration Materials: An Empirical Study Using Rasch Modeling. *Jurnal Pendidikan* <https://journal.unnes.ac.id/nju/JPF1/article/view/47493>
- Treagust, D. F. (2006). Diagnostic assessment in science as a means to improving teaching, learning and retention. *Proceedings of The Australian Conference on Science and Mathematics Education*.
- Vosniadou, S. (2019). The Development of Students' Understanding of Science. *Frontiers in Education*, 4. <https://doi.org/10.3389/educ.2019.00032>
- Zitha, I., Mokganya, M. G., & Manyage, T. (2023). Integration of Blended Learning in the Advent of COVID-19: Online Learning Experiences of the Science Foundation Students. *Citation: Zitha, I.* <https://doi.org/10.3390/educsci13070704>
- Zynuddin, S. N., & Sumintono, B. (2024). Longitudinal Study Of School Climate Instrument With Secondary School Students: Validity And Reliability Analysis With The Rasch Model. *Malaysian Online Journal of Educational Management*, 12(4), 24–41. <https://doi.org/10.22452/mojem.vol12no4.2>