



Assessing Assessment of Thermodynamics Concept Understanding: A Systematic Review

Delilah Nur Maisyaroh*, Sutopo, & Parno

Pendidikan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Malang, Indonesia

*Corresponding author: delilah.nur.2403218@students.um.ac.id

Abstract: *This study presents a systematic review of scholarly articles examining the development of thermodynamics concept assessment tools in physics education from 2010 to 2025. The review aims to analyze the characteristics, scope, and quality of existing instruments while offering guidance for future improvements. Eleven assessment tools were identified and reviewed using a structured protocol aligned with PRISMA and the guidelines of Petticrew & Roberts. The findings reveal a strong reliance on multiple-choice formats and a primary focus on fundamental thermodynamics concepts, chiefly heat, temperature, and the laws of thermodynamics. However, key conceptual areas such as entropy in open systems, real gas behavior, and statistical interpretations of thermodynamic laws are often overlooked. Furthermore, although many instruments report content validation by experts, only a few provide evidence of construct validity or internal consistency. The review recommends that future thermodynamics assessments align with students' cognitive development, maintain a balance between item depth and efficiency, and include underrepresented conceptual domains. Overall, the review advocates for the creation of more comprehensive and effective assessment tools in the field of thermodynamics.*

Keywords: *assessment, concept comprehension, instrument, physics education, thermodynamics*

Asesmen Pemahaman Konsep Termodinamika: Systematic Literature Review

Abstrak: Studi ini menyajikan tinjauan sistematis terhadap artikel-artikel penelitian yang berfokus pada pengembangan instrumen asesmen pemahaman konsep termodinamika dalam pendidikan fisika dari tahun 2010 hingga 2025. Tujuan dari tinjauan ini adalah untuk mengevaluasi karakteristik, cakupan, dan kualitas instrumen yang telah ada serta memberikan arahan bagi pengembangan di masa mendatang. Sebanyak 11 instrumen asesmen diidentifikasi dan dianalisis menggunakan protokol tinjauan terstruktur berdasarkan pedoman PRISMA dan Petticrew & Roberts. Tinjauan ini menunjukkan dominasi format pilihan ganda serta penekanan pada penilaian konsep-konsep inti termodinamika seperti kalor, suhu, dan hukum-hukum termodinamika. Namun demikian, beberapa area konseptual seperti entropi pada sistem terbuka, perilaku gas nyata, dan interpretasi statistik terhadap hukum-hukum termodinamika masih kurang terukur. Selain itu, ditemukan bahwa meskipun sebagian besar instrumen melaporkan validitas isi melalui penilaian ahli, hanya sedikit yang memberikan bukti validitas konstruk atau konsistensi internal. Disarankan agar asesmen termodinamika di masa depan mempertimbangkan keselarasan dengan tingkat kognitif siswa, keseimbangan antara kedalaman butir soal dan efisiensi, serta mencakup domain konsep yang selama ini terabaikan. Tinjauan ini mendukung pengembangan instrumen asesmen yang lebih efektif dan komprehensif dalam pendidikan termodinamika.

Kata kunci: asesmen, instrumen, pemahaman konsep, pendidikan fisika, termodinamika

PENDAHULUAN

Termodinamika merupakan topik yang kompleks dan multidimensional, sehingga menimbulkan tantangan bagi siswa di berbagai jenjang pendidikan. Termodinamika secara mendalam mengkaji energi, entropi, dan hubungan-hubungan di antara keduanya, serta menjadi dasar dalam menganalisis berbagai proses dan siklus termodinamika (Larsen et al., 2015; Li et al., 2020). Konsep-konsep inti dalam termodinamika, seperti energi dalam, entropi, kalor, dan kerja, bersifat konseptual dan tidak dapat diamati secara empiris, sehingga memerlukan keterampilan representasi mental yang kuat serta pemahaman mendalam terhadap hukum-hukum dasar yang mengaturnya (Brundage et al., 2025; Liu, 2022). Selain itu, perbedaan antara variabel-variabel termodinamika, seperti *state function* (seperti energi dalam, entropi, entalpi) dan *path function* (kalor dan kerja), sering kali menjadi sumber kebingungan bagi siswa (Brown & Singh, 2022). Sejumlah penelitian juga menyoroti berbagai kesulitan yang dihadapi siswa dalam mempelajari termodinamika, seperti kegagalan mengenali bahwa perubahan energi dalam bernilai nol pada gas ideal selama proses isothermal meskipun terjadi perubahan volume, kesalahpahaman bahwa kalor dan kerja selalu bernilai nol dalam semua proses siklik, atau anggapan bahwa entropi selalu meningkat bahkan dalam proses ireversibel (Brundage et al., 2025; Loverude, 2023). Konsepsi naif yang terus-menerus ditemukan menunjukkan pentingnya keberadaan asesmen yang dirancang dengan baik untuk secara akurat mendiagnosis dan mengatasi pemahaman konseptual siswa terhadap prinsip-prinsip termodinamika.

Eksplorasi pemahaman dan kesulitan siswa dalam mempelajari termodinamika mencakup penyelidikan terhadap penguasaan konsep siswa mengenai perpindahan kalor, kerja, energi dalam, proses gas ideal, entalpi, entropi, dan siklus mesin kalor (Brundage et al., 2025; Loverude, 2023; Smith et al., 2015). Pengembangan instrumen asesmen konseptual termodinamika dalam konteks pendidikan fisika diawali dengan HTCE (*Heat and Temperature Conceptual Evaluation*) yang dikembangkan oleh Thornton & Sokoloff (2001). Sejak saat itu, berbagai instrumen asesmen termodinamika telah dikembangkan, banyak di antaranya memuat butir soal yang secara khusus menguji Hukum I dan II Termodinamika. Instrumen-instrumen tersebut antara lain TCE (*Thermal Concept Evaluation*) (Yeo & Zadnik, 2001), TCS (*Thermodynamic Concept Survey*) (Wattanakasiwich et al., 2013), TCT (*Thermodynamic Concept Test*) (Kulkarni & Tambade, 2013), TDT (*Thermodynamic Diagnostic Test*) (Kamcharean & Wattanakasiwich, 2016), 4T-TDT (*Four-Tier Thermodynamic Diagnostic Test*) (Bhakti et al., 2022), HTCQ (*Heat and Temperature Concept Questionnaire*) (Kibirige, 2021), U-STEP (*Upperlevel Statistical and Thermodynamics Evaluation for Physics*) (Rainey et al., 2022), STPFaSL (*Survey of Thermodynamics Processes and First and Second Laws*) (Brown & Singh, 2021), and STPFaSL-Long (Brundage & Singh, 2023).

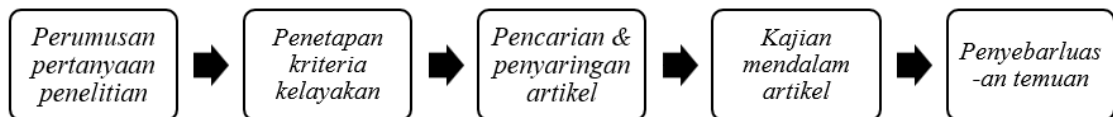
Kajian terhadap pengembangan instrumen asesmen pemahaman konseptual telah banyak dilakukan. Beberapa penelitian yang berfokus pada instrumen asesmen konsep sains menunjukkan bahwa format diagnostik yang umum digunakan meliputi wawancara, tes uraian terbuka maupun tertutup, tes pilihan ganda, serta tes pilihan ganda bertingkat (Gurel et al., 2015; Soeharto et al., 2019; Toma & Lederman, 2022). Sebagian besar studi tersebut mengkaji instrumen diagnostik sains secara umum, dengan menyoroti jenis format yang digunakan beserta kelebihan dan keterbatasannya, seperti permasalahan validitas (baik konten maupun konstruk) dan reliabilitas. Namun, belum banyak penelitian yang secara sistematis meninjau instrumen asesmen untuk konsep-konsep tertentu dalam fisika, khususnya konsep-konsep dalam termodinamika. Studi-studi yang ada umumnya hanya menitikberatkan pada aspek validitas dan reliabilitas, tanpa mengkaji cakupan dan

kekhususan konsep termodinamika, kesesuaian dengan tingkat pemahaman siswa, atau efektivitas dan efisiensi instrumen dalam mengevaluasi penguasaan konseptual di bidang termodinamika.

Tinjauan ini secara sistematis mengeksplorasi dan menganalisis 11 instrumen yang telah dikembangkan dan banyak digunakan untuk mengukur pemahaman konsep termodinamika selama 15 tahun terakhir, yaitu dari tahun 2010 hingga 2025. Hal ini agar kajian sistematis tetap relevan bagi kebaruan penelitian pada perkembangan instrument termodinamika bidang pendidikan fisika. Kajian ini menyajikan karakteristik dan kualitas instrumen-instrumen yang ada serta membahas hal-hal penting yang perlu dipertimbangkan dalam merancang asesmen pemahaman konsep termodinamika yang lebih efektif kedepannya. Kajian ini disusun secara sistematis dengan tujuan menyajikan secara menyeluruh dan terstruktur karakteristik serta kualitas instrumen asesmen pemahaman konsep termodinamika yang telah ada, sekaligus mengidentifikasi aspek-aspek kunci yang perlu diperhatikan dalam pengembangan instrumen asesmen konsep termodinamika ke depan.

METODE

Penelitian ini menggunakan pendekatan yang sistematis dan terstruktur berdasarkan pedoman praktis untuk tinjauan sistematis yang dikemukakan oleh Petticrew & Roberts (2006). Proses tinjauan dilakukan secara menyeluruh melalui serangkaian tahapan yang telah ditentukan dengan jelas, meliputi perumusan pertanyaan penelitian, penetapan kriteria kelayakan, pencarian dan penyaringan literatur secara menyeluruh, penilaian kritis terhadap studi yang terpilih, telaah mendalam, serta penyebaran temuan. Seluruh tahapan ini dilaksanakan secara berurutan untuk menjamin ketelitian metodologis dalam mengidentifikasi dan mensintesis literatur yang relevan, sebagaimana ditunjukkan pada Gambar 1.

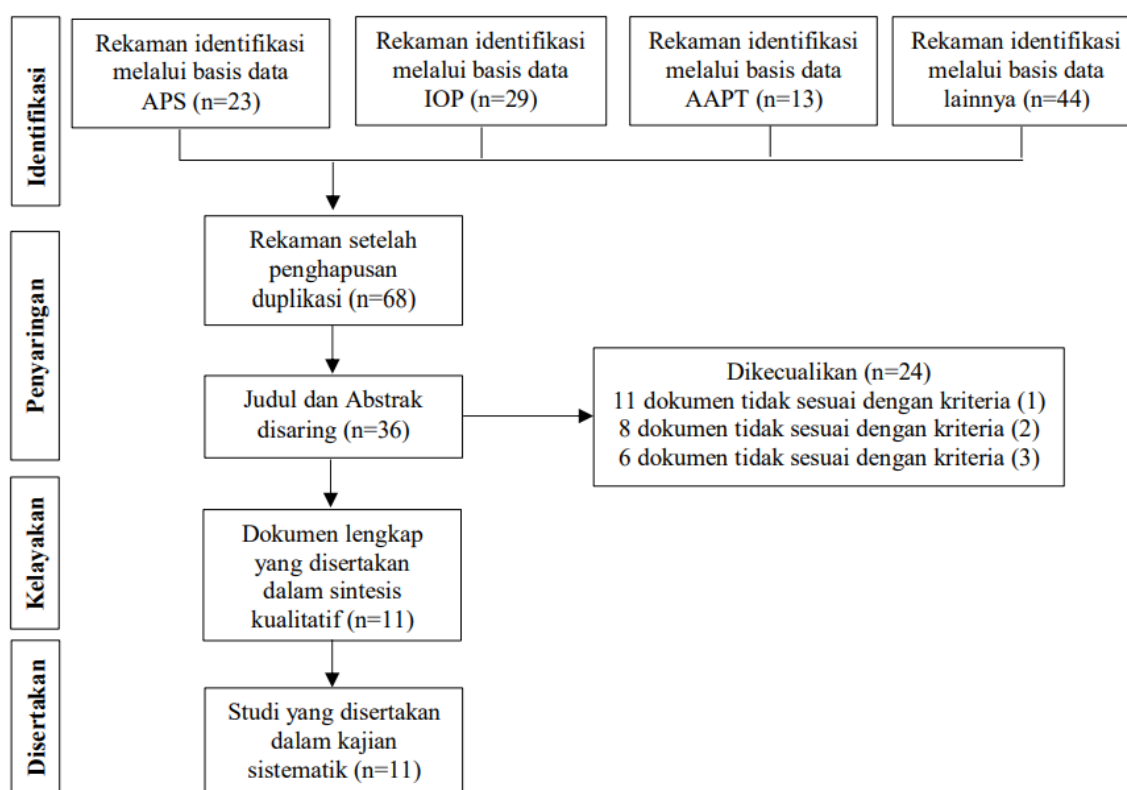


Gambar 1. Modifikasi Skema SLR (Petticrew & Roberts, 2006)

Penelitian ini dipandu oleh empat pertanyaan penelitian utama: (i) Apa saja karakteristik dan kualitas instrumen asesmen untuk konsep-konsep termodinamika?; (ii) Konsep-konsep termodinamika mana yang telah diases secara komprehensif?; (iii) Seberapa efektif dan efisien instrumen-instrumen tersebut dalam mengevaluasi pemahaman konseptual terhadap konsep-konsep termodinamika?; dan (iv) Arah pengembangan seperti apa yang sebaiknya diambil untuk pengembangan instrumen asesmen konsep termodinamika di masa depan?. Tahap selanjutnya dari tinjauan ini adalah menetapkan kriteria kelayakan untuk menentukan relevansi artikel yang akan dimasukkan dalam analisis. Suatu artikel dianggap relevan apabila memenuhi kriteria inklusi berikut: (1) artikel secara eksplisit membahas konsep-konsep termodinamika atau subtopiknya dalam bidang fisika; (2) artikel melaporkan dan mendiskusikan instrumen asesmen untuk konsep-konsep termodinamika; dan (3) artikel diterbitkan dalam rentang waktu antara tahun 2010 hingga 2025.

Proses peninjauan artikel dalam penelitian ini mengikuti diagram PRISMA (Liberati et al., 2009; Moher et al., 2009), sebagaimana ditunjukkan pada Gambar 2. Kerangka ini digunakan untuk mengatur langkah-langkah pencarian dan seleksi artikel secara sistematis.

Artikel yang relevan diidentifikasi dari jurnal-jurnal ilmiah di bidang pendidikan sains dan pendidikan fisika, serta melalui basis data akademik utama seperti IOP Publishing, APS, AAPT, dan sumber-sumber lainnya. Strategi pencarian menggunakan kombinasi kata kunci yang berkaitan dengan konsep-konsep termodinamika, termasuk energi dalam, kerja, perpindahan kalor, dan entropi. Selain itu, teknik *snowball* juga diterapkan dengan menelusuri daftar pustaka dari sejumlah artikel yang memenuhi syarat guna menemukan studi lain yang memiliki relevansi.



Gambar 2. Modifikasi Diagram Alur PRISMA

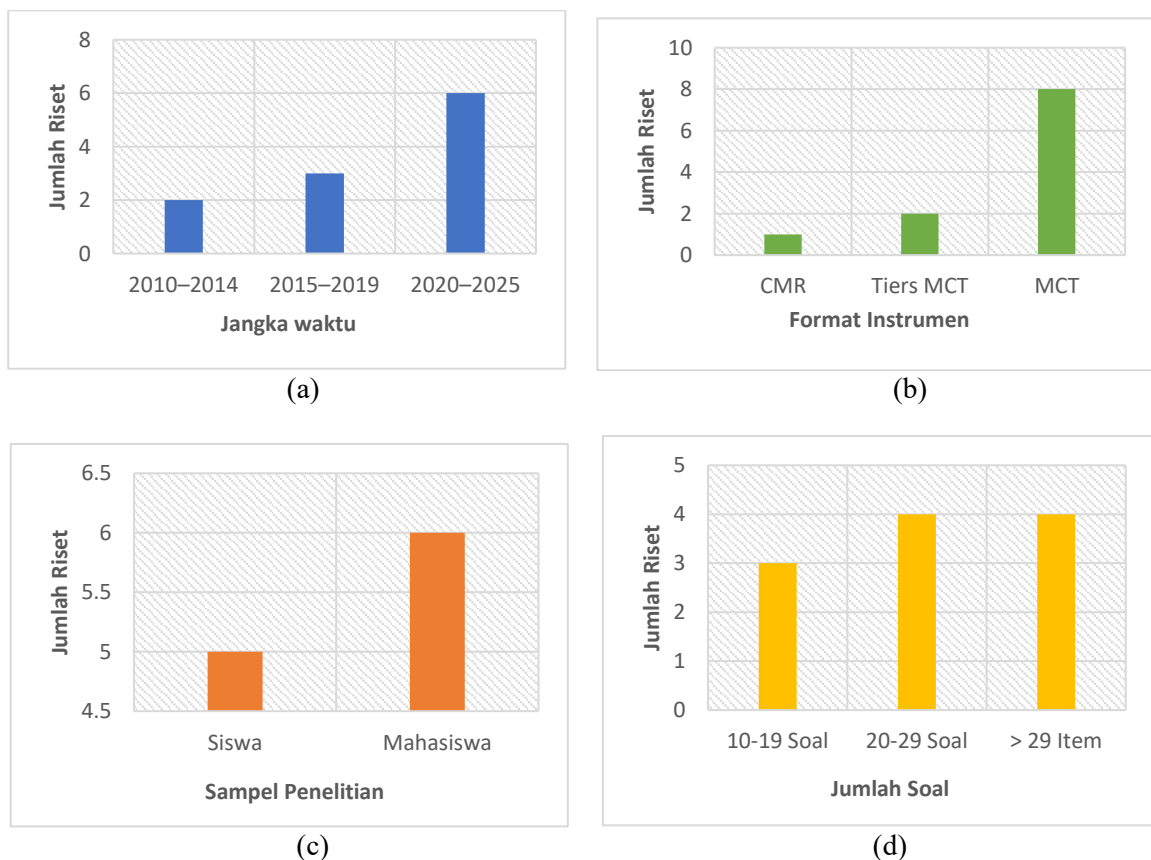
Proses penelusuran literatur yang dilakukan melalui sejumlah basis data utama dalam bidang pendidikan sains dan fisika menghasilkan 109 artikel. Dari jumlah tersebut, 41 artikel dieliminasi karena merupakan duplikasi. Selanjutnya, 68 artikel yang tersisa disaring berdasarkan judul dan abstrak, dengan mempertimbangkan relevansinya terhadap fokus penelitian. Hasil penyaringan awal tersebut menghasilkan 36 artikel yang dinilai berpotensi untuk dianalisis lebih lanjut, sementara artikel lainnya dievaluasi ulang menggunakan kriteria inklusi yang telah ditentukan sebelumnya. Akhirnya, 11 artikel dinyatakan memenuhi seluruh kriteria dan dipilih untuk ditelaah secara mendalam. Proses peninjauan ini mencakup evaluasi kritis dan sintesis terhadap hasil penelitian, dengan fokus pada karakteristik dan mutu instrumen asesmen konsep termodinamika. Penilaian dilakukan berdasarkan indikator umum seperti validitas, reliabilitas, cakupan materi, serta efektivitas dan efisiensi instrumen. Temuan utama dan rekomendasi dari analisis tersebut dipaparkan dalam bagian hasil tinjauan ini.

HASIL DAN PEMBAHASAN

Karakteristik dan Kualitas Instrumen Asesmen Termodinamika

Gambar 3 menyajikan karakteristik dari 11 instrumen asesmen konsep termodinamika yang telah dianalisis dan dilaporkan, baik dalam kaitannya dengan alat asesmen maupun pemahaman konseptual siswa terhadap termodinamika. Gambar tersebut menunjukkan adanya peningkatan yang signifikan dalam penelitian yang berfokus pada pengembangan instrumen-instrumen ini dalam beberapa tahun terakhir. Seperti yang ditunjukkan pada Gambar 3(a), tren yang meningkat sejak tahun 2010–2025 yang mencerminkan minat yang tinggi dikalangan peneliti untuk merancang asesmen yang bertujuan mendiagnosis pemahaman konseptual siswa, khususnya sebelum dan sesudah proses pembelajaran termodinamika (Brundage et al., 2024, 2025).

Jika dibandingkan antara periode 2020–2025 dan periode sebelumnya (2015–2019), terjadi peningkatan sebanyak dua kali lipat dalam jumlah studi pengembangan instrumen, yang menunjukkan lonjakan perhatian terhadap bidang ini. Seperti terlihat pada Gambar 3(c), sebagian besar studi menggunakan mahasiswa sebagai partisipan penelitian. Selain itu, sebagian besar instrumen asesmen konsep termodinamika dikembangkan dalam bentuk tes pilihan ganda atau MCT (*multiple choice test*) dan format tes bertingkat atau Tier-MCT (*multiple-tier test*), sebagaimana ditampilkan pada Gambar 3(b). Tes pilihan ganda banyak digunakan dalam pendidikan sains karena dianggap sebagai sarana yang efisien untuk mengeksplorasi, mendiagnosis, dan menilai pemahaman siswa terhadap konsep-konsep ilmiah tertentu (Mohd Dzin & Lay, 2021). Meskipun format ini mendominasi, masih terdapat potensi besar untuk mengeksplorasi format asesmen lainnya, seperti wawancara, tes esai, ataupun kombinasi antara tes tulis dan wawancara.



Gambar 3. Karakteristik Instrument Asesmen Termodinamika (a) Tahun Publikasi, (b) Format Instrumen, (c) Sampel Penelitian, (d) Jumlah Soal

Setiap format asesmen memiliki kelebihan dan keterbatasan tersendiri. Dari sebelas instrumen asesmen konsep termodinamika yang ditinjau, tidak ada yang menggunakan wawancara terstruktur maupun semi-terstruktur sebagai bentuk utama asesmen. Wawancara umumnya digunakan dalam tahap pengembangan atau validasi, karena mampu mengungkap struktur kognitif siswa secara mendalam, mengidentifikasi miskonsepsi, serta menilai pemahaman konseptual secara rinci (Gurel et al., 2015; Kaltakci-Gurel et al., 2017; Young et al., 2018). Sebagai contoh, pengembangan TCS menggunakan wawancara di tahap awal untuk menemukan miskonsepsi umum. Sementara itu, STPFaLS, STPFaLS-Long, dan U-STEP memanfaatkan wawancara *think-aloud* untuk memvalidasi butir soal, mengukur tingkat pemahaman siswa, dan menyempurnakan versi akhir instrumen. Pada TDT, wawancara digunakan untuk menyusun distraktor dalam format dua tingkat (*two tiers*). Meskipun informatif, wawancara memerlukan waktu, keahlian komunikasi, dan analisis data berupa koding yang cukup kompleks (Langley & Meziani, 2020), serta kurang efisien untuk sampel besar. Demikian pula, soal uraian yang memberikan kebebasan bagi siswa untuk mengemukakan alasan dan membangun argumen, sehingga menawarkan eksplorasi yang lebih dalam terhadap pemahaman konseptual mereka (Melovitz Vasan et al., 2018). Namun, respons soal uraian seringkali sulit dianalisis, terutama ketika siswa memberikan jawaban yang samar atau kurang terstruktur (He & Schonlau, 2020).

Tes dengan format pilihan ganda (MCT) telah menjadi bentuk evaluasi yang paling banyak diterapkan, sebagaimana terlihat pada TCE, TCS, TCT, HTCE, HTCQ, dan STPFaLS. Format ini unggul dalam hal efisiensi dan kemampuan penilaian skala besar (Mohd Dzin & Lay, 2021). Namun, pengembangannya tidak sederhana; misalnya, instrumen STPFaLS-Long, yang merupakan survei termodinamika terbesar saat ini, membutuhkan riset dan revisi iteratif secara berkala (Brundage et al., 2024, 2025; Brundage & Singh, 2023). Kendati demikian, format pilihan ganda atau MCT memiliki kelemahan dalam menggambarkan kedalaman pemahaman siswa serta membedakan antara jawaban yang benar karena pemahaman ilmiah atau sekadar menebak (Izci & Caliskan, 2017). Karenanya penggunaan diagnostic tes bertingkat seperti TDT dan 4T-TDT juga dikembangkan untuk mendiagnosis miskonsepsi siswa (Bhakti et al., 2022; Kamcharean & Wattanakasiwich, 2016). Lebih lanjut, penggabungan komponen wawancara dalam proses pengembangan MCT, seperti dilakukan dalam TCS, TDT, U-STEP, dan STPFaLS, menjadi strategi yang direkomendasikan untuk memperkaya asesmen dan meningkatkan validitas hasilnya (Brundage et al., 2025; Langley & Meziani, 2020; Rainey et al., 2022).

Validitas dan Reliabilitas Instrumen

Evaluasi terhadap instrumen asesmen konsep termodinamika umumnya mengacu pada standar validitas dan reliabilitas yang diakui secara internasional. Aspek yang dianalisis mencakup validitas isi, konstruk, prediktif, diskriminatif, serta reliabilitas internal terkait konsistensi dan stabilitas hasil asesmen. Namun, dari seluruh instrumen yang dikaji, tidak ada yang melaporkan validitas prediktif maupun diskriminatif, beberapa bahkan tidak menyebutkan data validitas atau reliabilitasnya. Sebagian besar instrumen melaporkan validitas konten, biasanya diperoleh melalui penilaian para ahli terhadap kesesuaian butir soal dengan karakteristik peserta dan cakupan materi. Namun, validitas konstruk dan reliabilitas internal belum secara konsisten dilaporkan. Studi lanjutan terhadap instrumen-instrumen ini telah menggunakan pendekatan seperti *Classical Test Theory* (CTT), *Exploratory Factor Analysis* (EFA), dan *Differential Item Functioning* (DIF) untuk memperkuat bukti validitas dan reliabilitas.

Validitas konten menjadi aspek krusial dalam pengembangan instrumen karena menilai sejauh mana item soal relevan, mencerminkan cakupan konsep, serta dapat dipahami oleh peserta tes (Almanasreh et al., 2019; Toma & Lederman, 2022). Proses ini umumnya melibatkan ahli yang meninjau kejelasan, tingkat kesulitan, kesesuaian materi, dan potensi kesalahan pada butir soal (Connell et al., 2018), sehingga memperkuat keandalan instrumen dalam mengukur pemahaman konsep yang akan diuji (Beck, 2020).

Tabel 1. Validitas dan Reliabilitas Intrumen Asesmen Termodinamika

Instrumen	Riset	Validitas Konten	Validitas Konstruk	Reliabilitas
TCE	(Stylos et al., 2021)	A,E	IRT	KR-20
TCS	(Wattanakasiwich et al., 2013)	A,E	KGV	KR-20
TCS	(Malik et al., 2019)	A		
TCT	(Kulkarni & Tambade, 2013)	A,E	CTT	KR-21
TDT	(Kamcharean & Wattanakasiwich, 2016)	A,E	Wawancara	KR-20
4D-TDT	(Bhakti et al., 2022)	A		
HTCE	(Nugraha et al., 2017)	A	Analisis Konsep	
HTCQ	(Kibirige, 2021)	A,E		Multi-metric
U-STEP	(Rainey et al., 2022)	A,E	CTT, EFA, DIF	CR-A
STPFaSL	(Brown & Singh, 2021)	A,E	CTT, Benchmarking	KR-20
STPFaSL-Long	(Brundage & Singh, 2023)	A,E	CTT, Correlational	KR-20

A : Validasi Ahli; E : Validasi Empirik; KR-20 : Kuder & Richardson; CR-A: Cronbach's Alpha; IRT : *Item Response Theory*; KGV : *Known Group Validity*

Konsep Termodinamika yang Telah Diuji

Pemahaman konseptual merujuk pada penguasaan pengetahuan yang bersifat abstrak dan menyeluruh terhadap konsep-konsep inti maupun pendukung dalam suatu bidang ilmu (Harjuna et al., 2024; Schneider & Stern, 2010). Membangun pemahaman siswa terhadap konsep-konsep fundamental menjadi tujuan sentral dalam pengajaran fisika (Xie et al., 2021; Xu et al., 2020). Pemanfaatan strategi pembelajaran inovatif yang mengintegrasikan aktivitas membaca, pemetaan konsep, dan berbagi pengetahuan terbukti mampu meningkatkan hasil belajar siswa serta memperkuat pemahaman konseptual (Widyaningsih & Yusuf, 2019). Bagi pendidik, pemahaman terhadap konsep awal yang dimiliki siswa, terutama ketika mencakup miskonsepsi atau pengetahuan awal yang tidak sesuai dengan pandangan ilmiah, sangatlah penting (Bozzi et al., 2019). Identifikasi terhadap konsepsi naif ini menjadi dasar yang berharga dalam merancang strategi pendekatan model pembelajaran yang edukatif yang efektif dalam mengembangkan pemahaman konsep siswa (Resbiantoro & Setiani, 2022).

Instrumen asesmen yang dirancang untuk mengidentifikasi tingkat pemahaman konseptual siswa dalam topik termodinamika umumnya dapat dikelompokkan ke dalam tiga kategori utama. Pertama, instrumen yang dirancang khusus untuk menilai pemahaman siswa mengenai konsep panas dan suhu, seperti HTCE, HTCQ, dan TCE. Kedua, instrumen yang menargetkan pemahaman terhadap proses dan hukum-hukum

termodinamika, meliputi TCS, STPFaLS, STPFaLS-Long, U-STEP, dan TDT. Ketiga, instrumen yang difokuskan pada identifikasi miskonsepsi siswa terhadap konsep-konsep termodinamika, seperti 4T-TDT, TDT, TCS (Gender), STPFaLS, dan U-STEP.

Rangkuman konsep-konsep termodinamika yang diukur oleh masing-masing instrumen disajikan pada Tabel 2. Secara keseluruhan, instrumen-instrumen tersebut mencakup berbagai topik inti dalam termodinamika, seperti panas, suhu, kesetimbangan termal, perpindahan kalor, kapasitas kalor, perubahan fasa, dan hukum gas ideal. Selain itu, sebagian besar juga mengevaluasi pemahaman siswa terhadap hukum termodinamika nol, pertama, dan kedua, serta konsep-konsep dasar seperti usaha, energi, dan perubahan keadaan sistem.

Tabel 2. Konsep Termodinamika yang Telah Diuji

Instrumen	Format	Butir Soal	Konsep Termodinamika
TCE	Pilihan Ganda (MCT)	26	Konsepsi alternatif terkait suhu, perpindahan kalor, dan karakteristik termal
TCS	Pilihan Ganda (MCT)	35	Suhu, perpindahan kalor, hukum gas ideal, dan hukum pertama termodinamika
TCS	Pilihan Ganda (MCT)	35	Fokus pada diagnosis berbasis gender terhadap hukum pertama dan konsep termodinamika
TCT	Pilihan Ganda (MCT)	24	Hukum-hukum termodinamika, konsep kalor, persamaan keadaan, proses isothermal, isokhorik, adiabatik, dan mesin kalor
TDT	Two-Tier MCT	15	Miskonsepsi umum tentang hukum nol, pertama, dan kedua termodinamika
4D-TDT	Four-Tier MCT	20	Diagnosis mendalam terhadap miskonsepsi pada prinsip dan hukum termodinamika
HTCE	Pilihan Ganda (MCT)	27	Suhu, perpindahan kalor, laju kalor, pendinginan, kalorimetri, persepsi inderawi terhadap panas, kalor jenis, perubahan fase, dan konduktivitas termal
HTCQ	Pilihan Ganda (MCT)	11	Kalor jenis, perubahan fase, dan konduktivitas termal
U-STEP	(MCR)	15	Energi, entropi, dan interpretasi statistik dalam sistem termodinamika
STPFaSL	Pilihan Ganda (MCT)	33	Proses termodinamika, transfer energi, mesin kalor, dan hukum-hukum termodinamika
STPFaSL-Long	Pilihan Ganda (MCT)	45	Konsep perluasan: entropi, kerja, konservasi energi, dan pemodelan sistem tingkat lanjut

MCR : *Multiple Coupled Response*

Sejumlah konsep penting dalam termodinamika masih kurang terwakili dalam instrumen asesmen yang ada saat ini. Konsep-konsep tersebut meliputi topik yang lebih lanjut atau bersifat mikroskopik, seperti entropi pada sistem terbuka, perilaku gas non-ideal (gas nyata), serta interpretasi statistik terhadap entropi dan ketidakteraturan. Selain itu, area penting lainnya yang jarang disentuh mencakup energi bebas Gibbs, entalpi, dan potensial termodinamika, yang semuanya sangat krusial untuk menjembatani hukum-

hukum makroskopik dengan perilaku mikroskopik. Kesenjangan ini menunjukkan perlunya pengembangan instrumen asesmen di masa mendatang yang mampu mengintegrasikan prinsip-prinsip termodinamika yang lebih mendalam dan abstrak, terutama yang mendukung transisi konseptual siswa dari termodinamika klasik menuju termodinamika statistik.

Efektivitas dan Efisiensi Instrumen

Berdasarkan jumlah butir soal yang ditampilkan dalam Gambar 3(d), instrumen asesmen konsep termodinamika yang dikembangkan dapat dikelompokkan ke dalam tiga kategori, yaitu asesmen dengan 10–19 butir soal, 20–29 butir soal, dan lebih dari 29 butir soal. Beberapa instrumen seperti *TCS*, *STPFaLS*, dan *STPFaLS-Long* termasuk dalam kategori dengan jumlah soal lebih dari 29. Instrumen dengan jumlah soal yang banyak menuntut siswa untuk mengerahkan waktu, upaya kognitif, dan konsentrasi yang tinggi dalam waktu yang lama, sehingga berpotensi menimbulkan kelelahan mental. Bagi pendidik, asesmen yang panjang juga memerlukan waktu lebih untuk dianalisis dan diinterpretasikan, sehingga kurang ideal untuk evaluasi cepat dan praktis. Sebaliknya, instrumen semacam ini lebih cocok digunakan dalam kajian mendalam untuk mengungkap tingkat pemahaman konseptual siswa secara lebih komprehensif. Di sisi lain, instrumen yang terdiri dari terlalu sedikit soal berisiko tidak mampu mencakup seluruh ruang lingkup konsep-konsep termodinamika yang penting. Hal tersebut mengisyaratkan pengembangan instrumen di masa mendatang sebaiknya mempertimbangkan jumlah soal yang optimal cukup untuk menilai pemahaman siswa secara menyeluruh, namun tetap ringkas agar efisien dalam pelaksanaan (Boone & Scantlebury, 2006). Jumlah butir soal yang ideal akan mendukung terciptanya alat asesmen yang efektif dan praktis, serta mampu mengukur pemahaman konseptual dan tingkat kompetensi siswa dalam termodinamika secara akurat

Implikasi untuk Pengembangan Instrumen Termodinamika Masa Mendatang

Penelitian dan penyempurnaan yang berkelanjutan sangat penting untuk meningkatkan pemahaman konseptual siswa dalam bidang termodinamika. Meskipun berbagai instrumen asesmen konsep termodinamika sebelumnya telah memberikan kontribusi yang kuat, pengembangan instrumen di masa depan perlu diarahkan untuk menjadi lebih kuat, terfokus, dan relevan secara kontekstual. Salah satu aspek krusial dalam pengembangan tersebut adalah penjaminan bukti validitas dan reliabilitas yang kuat, termasuk validitas isi, validitas konstruk, validitas prediktif, serta konsistensi internal dan stabilitas waktu. Validitas isi harus diuji melalui evaluasi para ahli untuk memastikan keterkaitan butir soal dengan konsep termodinamika yang dituju, sementara validitas konstruk perlu dibuktikan melalui uji coba terhadap sampel siswa untuk mengevaluasi sejauh mana siswa dapat memahami dan merespons butir instrumen dengan benar. Selain itu, cakupan dan kedalaman konsep-konsep termodinamika yang diukur juga perlu menjadi perhatian utama. Instrumen yang ada saat ini, seperti *TCE*, *HTCE*, dan *HTCQ*, banyak berfokus pada konsep dasar seperti panas dan suhu, sementara instrumen seperti *TCS*, *STPFaLS*, *U-STEP*, dan *TDT* menjangkau konsep yang lebih luas seperti hukum termodinamika, entropi, dan perpindahan energi. Namun demikian, masih terdapat sejumlah konsep penting yang kurang terakomodasi dalam asesmen saat ini, misalnya entropi dalam sistem terbuka, perilaku gas nyata (non-ideal), dan interpretasi mikroskopik terhadap hukum-hukum termodinamika. Oleh karena itu, asesmen ke depan sebaiknya mencakup

pengukuran yang komprehensif namun tetap terfokus, agar sejalan dengan kompleksitas materi termodinamika.

Keselarasannya dengan jenjang pendidikan siswa juga menjadi faktor penting (Harjuna et al., 2024). Materi termodinamika diajarkan mulai dari tingkat sekolah menengah hingga perguruan tinggi, dan tuntutan kognitif berbeda pada masing-masing jenjang. Oleh sebab itu, instrumen asesmen perlu dikembangkan secara spesifik untuk siswa SMP, SMA, atau mahasiswa, agar tingkat kesulitan dan konteks soal sesuai dengan karakteristik peserta didik. Hal ini juga berarti melibatkan ahli sesuai jenjang saat proses pengembangan, guna menghindari ketidaksesuaian seperti soal yang terlalu rumit untuk siswa SMP, atau sebaliknya, terlalu sederhana untuk mahasiswa. Contohnya, TDT dan 4T-TDT dikembangkan untuk siswa SMA, sedangkan STPFaLS dan U-STEP lebih cocok untuk diagnosis di tingkat universitas. Terakhir, aspek efisiensi dan efektivitas juga harus menjadi fokus utama dalam perancangan asesmen mendatang. Instrumen dengan jumlah soal terlalu banyak, seperti STPFaLS-Long, memang memberikan wawasan yang dalam, tetapi menyita waktu dan menguras energi, baik bagi siswa maupun pengajar. Idealnya, instrumen asesmen konsep termodinamika di masa depan dirancang tidak melebihi 30 butir soal, dengan kualitas soal yang baik dan representatif. Jumlah tersebut dinilai cukup untuk memberikan data yang valid dan komprehensif tanpa membebani peserta tes. Hal ini juga akan memudahkan proses penilaian dan interpretasi hasil oleh guru atau dosen, sehingga umpan balik dapat diberikan secara lebih cepat dan efektif. Dengan demikian, riset mendatang terkait instrumen asesmen konsep termodinamika perlu menjembatani kebutuhan antara kedalaman materi dan efisiensi pelaksanaan, serta memastikan validitas dan kemudahan penggunaan di berbagai tingkat pendidikan.

SIMPULAN DAN SARAN

Tinjauan terhadap instrumen asesmen konsep termodinamika yang telah ada menunjukkan kemajuan signifikan dalam mengevaluasi pemahaman konseptual siswa, terutama pada topik-topik seperti kalor, suhu, dan hukum-hukum termodinamika. Namun demikian, masih terdapat sejumlah celah yang perlu diperhatikan, seperti keterbatasan dalam mengukur konsep termodinamika tingkat lanjut dan mikroskopis, ketidaksesuaian dengan jenjang pendidikan yang berbeda, serta tantangan terkait efisiensi instrumen dan pelaporan validitasnya. Pengembangan instrumen di masa mendatang perlu difokuskan pada penciptaan alat asesmen yang mencakup konsep secara menyeluruh, sesuai dengan tingkat pendidikan, memiliki validitas psikometrik yang kuat, serta efisien secara praktis, sehingga dapat membantu pendidik dalam mendiagnosis, mendukung, dan meningkatkan pemahaman siswa terhadap prinsip-prinsip termodinamika secara lebih efektif.

DAFTAR PUSTAKA

- Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of Methods Used for Estimating Content Validity. *Research in Social and Administrative Pharmacy*, 15(2), 214–221. <https://doi.org/10.1016/j.sapharm.2018.03.066>
- Beck, K. (2020). Ensuring Content Validity of Psychological and Educational Tests—The Role of Experts. *Frontline Learning Research*, 8(6), 1–37.
- Bhakti, Y. B., Astuti, I. A. D., & Prasetya, R. (2022). Four-Tier Thermodynamics Diagnostic Test (4T-TDT) to Identify Student Misconception. *KnE Social Sciences*, 106–116. <https://doi.org/10.18502/kss.v7i14.11958>
- Boone, W. J., & Scantlebury, K. (2006). The Role of Rasch Analysis When Conducting Science Education Research Utilizing Multiple-Choice Tests. *Science Education*, 90(2), 253–269. <https://doi.org/10.1002/sce.20106>

- Bozzi, M., Ghislandi, P., Tsukagoshi, K., Matsukawa, M., Wada, M., Nagaoka, N., Pnev, A. B., Zhirnov, A. A., Guillaume, G., & Zani, M. (2019). Highlight Misconceptions in Physics: A TIME project. *INTED2019 Proceedings*, 2520–2525. <https://doi.org/10.21125/inted.2019.0689>
- Brown, B., & Singh, C. (2021). Development and Validation of a Conceptual Survey Instrument to Evaluate Students' Understanding of Thermodynamics. *Physical Review Physics Education Research*, 17(1), 010104. <https://doi.org/10.1103/PhysRevPhysEducRes.17.010104>
- Brown, B., & Singh, C. (2022). Student Understanding of Thermodynamic Processes, Variables and Systems. *European Journal of Physics*, 43(5), 055705. <https://doi.org/10.1088/1361-6404/ac7af2>
- Brundage, M. J., Meltzer, D. E., & Singh, C. (2024). Investigating Introductory and Advanced Students' Difficulties with Entropy and the Second Law of Thermodynamics Using a Validated Instrument. *Physical Review Physics Education Research*, 20(2), 020110. <https://doi.org/10.1103/PhysRevPhysEducRes.20.020110>
- Brundage, M. J., Meltzer, D. E., & Singh, C. (2025). Investigating the Impact of Problem Properties on Introductory and Advanced Student Responses to Introductory Thermodynamics Conceptual Problems. *Physical Review Physics Education Research*, 21(1), 010127. <https://doi.org/10.1103/PhysRevPhysEducRes.21.010127>
- Brundage, M. J., & Singh, C. (2023). Development and Validation of a Conceptual Multiple-Choice Survey Instrument to Assess Student Understanding of Introductory Thermodynamics. *Physical Review Physics Education Research*, 19(2), 020112. <https://doi.org/10.1103/PhysRevPhysEducRes.19.020112>
- Connell, J., Carlton, J., Grundy, A., Taylor Buck, E., Keetharuth, A. D., Ricketts, T., Barkham, M., Robotham, D., Rose, D., & Brazier, J. (2018). The Importance of Content and Face Validity in Instrument Development: Lessons Learnt From Service Users when Developing the Recovering Quality of Life measure (ReQoL). *Quality of Life Research*, 27(7), 1893–1902. <https://doi.org/10.1007/s11136-018-1847-y>
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). *A Review and Comparison of Diagnostic Instruments to Identify Students' Misconceptions in Science*. <http://dx.doi.org/10.12973/eurasia.2015.1369a>
- Harjuna, R., Sutopo, S., & Munfaridah, N. (2024). Learning Based on Central Idea on the Topic of Electrical Circuits to Improve Students' Concept Understanding. *Kasuari: Physics Education Journal (KPEJ)*, 7(2), 446–456. <https://doi.org/10.37891/kpej.v7i2.781>
- He, Z., & Schonlau, M. (2020). Automatic Coding of Text Answers to Open-Ended Questions: Should You Double-Code the Training Data? *Social Science Computer Review*, 38(6), 754–765. <https://doi.org/10.1177/0894439319846622>
- 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 & Technological Education*, 35(2), 238–260. <https://doi.org/10.1080/02635143.2017.1310094>
- Kamcharean, C., & Wattanakasiwich, P. (2016). Development and Application of Thermodynamics Diagnostic Test to Survey Students' Understanding in Thermal Physics. *International Journal of Innovation in Science and Mathematics Education*, 24(2). <https://openjournals.library.sydney.edu.au/CAL/article/view/8710>
- Kibirige, I. (2021). Exploring the Prevalence of Misconceptions Regarding Heat and Temperature among Grade Nine Natural Science Learners. *Unnes Science Education Journal*, 10(3), 115–123. <http://dx.doi.org/10.15294/usej.v10i3.47258>

- Kulkarni, V. D., & Tambade, P. S. (2013). Assessing the Conceptual Understanding About Heat and Thermodynamics at Undergraduate Level. *European Journal Of Physics Education*, 4(2), 9–16. <https://doi.org/10.20308/EJPE.06598>
- Langley, A., & Meziani, N. (2020). Making Interviews Meaningful. *The Journal of Applied Behavioral Science*, 56(3), 370–391. <https://doi.org/10.1177/0021886320937818>
- Larsen, K., Hossain, A., & Weiser, M. (2015). Optimizing the Curriculum in Thermodynamics and Heat Transfer with Better Labs, In-Class Demonstrations and Interesting Realistic Problems to Enhance Learning. *ASME International Mechanical Engineering Congress and Exposition*, 57427, V005T05A028. <http://dx.doi.org/10.1115/IMECE2015-52264>
- Li, C., Li, B., Huang, J., & Li, C. (2020). Energy and Exergy Analyses of a Combined Infrared Radiation-Counterflow Circulation (IRCC) Corn Dryer. *Applied Sciences*, 10(18), 6289. <https://doi.org/10.3390/app10186289>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies that Evaluate Healthcare Interventions: Explanation and Elaboration. *Bmj*, 339. <https://doi.org/10.1136/bmj.b2700>
- Liu, J. Z. (2022). *Misconceptions in Thermodynamics*. Stanford University. Archived (PDF). <http://xenon.stanford.edu/~zjl/pdf/thermal.pdf>
- Loverude, M. E. (2023). Student Understanding of Thermal Physics. *The International Handbook of Physics Education Research: Learning Physics*, 3–1. https://doi.org/10.1063/9780735425477_003
- Malik, U., Angstmann, E. J., & Wilson, K. (2019). Learning and Conceptual Change in Thermal Physics Concepts: An Examination by Gender. *International Journal of Innovation in Science and Mathematics Education*, 27(1). <http://dx.doi.org/10.30722/IJISME.27.01.003>
- Melovitz Vasan, C. A., DeFouw, D. O., Holland, B. K., & Vasan, N. S. (2018). Analysis of Testing with Multiple Choice Versus Open-Ended Questions: Outcome-Based Observations in an Anatomy Course. *Anatomical Sciences Education*, 11(3), 254–261. <https://doi.org/10.1002/ase.1739>
- Mohd Dzin, N. H., & Lay, Y. F. (2021). Assessing the Validity and Reliability of Science Multiple Choice Test Using Rasch Dichotomous Measurement Model. *Journal of Baltic Science Education*, 20(6), 927–941. <https://doi.org/10.33225/jbse/21.20.927>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Bmj*, 339. <https://doi.org/10.1136/bmj.b2535>
- Nugraha, D. A., Suparmi, A., & Winarni, R. (2017). Students' Profile of Heat and Temperature Using HTCE in Undergraduate Physics. *Journal of Physics: Conference Series*, 909(1), 012059. <https://doi.org/10.1088/1742-6596/909/1/012059>
- Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide*. <https://doi.org/10.1002/9780470754887>
- Rainey, K. D., Vignal, M., & Wilcox, B. R. (2022). Validation of a Coupled, Multiple Response Assessment for Upper-Division Thermal Physics. *Physical Review Physics Education Research*, 18(2), 020116. <https://doi.org/10.1103/PhysRevPhysEducRes.18.020116>

- Resbiantoro, G., & Setiani, R. (2022). A Review of Misconception in Physics: The Diagnosis, Causes, and Remediation. *Journal of Turkish Science Education*, 19(2), 403–427. <http://dx.doi.org/10.36681/tused.2022.128>
- Schneider, M., & Stern, E. (2010). The Developmental Relations Between Conceptual and Procedural Knowledge: A Multimethod Approach. *Developmental Psychology*, 46(1), 178. <https://doi.org/10.1037/a0016701>
- Smith, T. I., Christensen, W. M., Mountcastle, D. B., & Thompson, J. R. (2015). Identifying Student Difficulties with Entropy, Heat Engines, and the Carnot Cycle. *Physical Review Special Topics - Physics Education Research*, 11(2), 020116. <https://doi.org/10.1103/PhysRevSTPER.11.020116>
- Soeharto, S., 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>
- Stylos, G., Sargioti, A., Mavridis, D., & Kotsis, K. T. (2021). Validation of the Thermal Concept Evaluation Test for Greek University Students' Misconceptions of Thermal Concepts. *International Journal of Science Education*, 43(2), 247–273. <https://doi.org/10.1080/09500693.2020.1865587>
- Thornton, R., & Sokoloff, D. (2001). Heat and Temperature Conceptual Evaluation (HTCE). *PhysPort Supporting Physics Teaching with Research-Based Resources*. <https://www.physport.org/assessments/assessment.cfm?I=16>
- Toma, R. B., & Lederman, N. G. (2022). A Comprehensive Review of Instruments Measuring Attitudes Toward Science. *Research in Science Education*, 52(2), 567–582. <http://dx.doi.org/10.1007/s11165-020-09967-1>
- Wattanawasiwich, P., Taleab, P., (2013). Construction and Implementation of a Conceptual Survey in Thermodynamics. *International*, Query date: 2025-04-23 19:24:01. <https://openjournals.library.sydney.edu.au/cal/article/view/6459>
- Widyaningsih, S. W., & Yusuf, I. (2019, October). Effect of RMS models (reading, mind mapping and sharing) based on e-learning to the learning outcomes of students in the teaching and learning subject. In *Journal of Physics: Conference Series* (Vol. 1321, No. 3, p. 032062). IOP Publishing. <http://doi.org/10.1088/1742-6596/1321/3/032062>
- Xie, L., Liu, Q., Lu, H., Wang, Q., Han, J., Feng, X., & Bao, L. (2021). Student Knowledge Integration in Learning Mechanical Wave Propagation. *Physical Review Physics Education Research*, 17(2), 020122. <https://doi.org/10.1103/PhysRevPhysEducRes.17.020122>
- Xu, W., Liu, Q., Koenig, K., Fritchman, J., Han, J., Pan, S., & Bao, L. (2020). Assessment of Knowledge Integration in Student Learning of Momentum. *Physical Review Physics Education Research*, 16(1), 010130. <https://doi.org/10.1103/PhysRevPhysEducRes.16.010130>
- Yeo, S., & Zadnik, M. (2001). Introductory Thermal Concept Evaluation: Assessing Students' Understanding. *The Physics Teacher*, Query date: 2025-04-23 19:24:01. <http://dx.doi.org/10.1119/1.1424603>
- Young, J. C., Rose, D. C., Mumby, H. S., Benitez-Capistros, F., Derrick, C. J., Finch, T., Garcia, C., Home, C., Marwaha, E., Morgans, C., Parkinson, S., Shah, J., Wilson, K. A., & Mukherjee, N. (2018). A Methodological Guide to Using and Reporting on Interviews in Conservation Science Research. *Methods in Ecology and Evolution*, 9(1), 10–19. <https://doi.org/10.1111/2041-210X.12828>