



PENGEMBANGAN INSTRUMEN PENILAIAN HIGHER ORDER THINKING SKILLS (HOTS) PADA MATERI SISTEM PENGAPIAN ELEKTRONIK SEPEDA MOTOR

Riyadi¹, Imam Mahir², M Rudyansyah³, Agung Gumelar⁴
Universitas Negeri Jakarta
Program Studi Pendidikan Teknik Mesin, Fakultas Teknik
e-mail: riyadi@unj.ac.id



ABSTRAK

Soal ujian pada mata pelajaran pemeliharaan kelistrikan sepeda motor di Sekolah Menengah Kejuruan (SMK) masih didominasi butir yang menguji ingatan, sehingga kemampuan siswa menganalisis dan mengevaluasi gangguan sistem pengapian jarang terukur. Penelitian ini bertujuan menghasilkan instrumen penilaian keterampilan berpikir tingkat tinggi (*Higher Order Thinking Skills*) pada materi sistem pengapian elektronik sepeda motor dan menguji kelayakannya. Penelitian pengembangan ini mengikuti model 4D (*define, design, develop, disseminate*) dari Thiagarajan. Lima belas butir pilihan ganda pada level menganalisis (C4) dan mengevaluasi (C5) divalidasi oleh dua ahli, lalu diujicobakan kepada 44 siswa kelas XII Program Keahlian Teknik dan Bisnis Sepeda Motor di dua SMK negeri di Jakarta. Data dianalisis dengan korelasi *point biserial*, rumus Kuder-Richardson 20 (KR-20), indeks daya pembeda, dan indeks tingkat kesukaran. Sebanyak 13 dari 15 butir dinyatakan valid ($r \text{ hitung} \geq 0,297$) dengan koefisien reliabilitas 0,71 (kategori tinggi). Di antara butir yang valid, sembilan butir memiliki daya pembeda baik, satu butir sangat baik, dan tiga butir cukup; enam butir tergolong sedang dan tujuh butir sukar, tanpa butir mudah. Ketiga belas butir tersebut dapat dipakai guru untuk mengukur kemampuan siswa mendiagnosis gangguan pengapian elektronik, sedangkan dua butir yang gugur perlu ditulis ulang sebelum digunakan.

Kata kunci: instrumen penilaian, berpikir tingkat tinggi, sistem pengapian elektronik, model 4D, sekolah menengah kejuruan

ABSTRACT

Test items in the motorcycle electrical maintenance subject at Indonesian vocational high schools still focus on recall, so students' ability to analyse and evaluate ignition faults is rarely measured. This study aimed to develop a Higher Order Thinking Skills (HOTS) assessment instrument for the motorcycle electronic ignition system and to examine its quality. We followed the 4D development model (*define, design, develop, disseminate*) proposed by Thiagarajan. Fifteen multiple-choice items at the analysing (C4) and evaluating (C5) levels were reviewed by two experts and then administered to 44 twelfth-grade students of the Motorcycle Engineering and Business programme at two public vocational high schools in Jakarta. Data were analysed using point-biserial correlation, the Kuder-Richardson 20 (KR-20) formula, the discrimination index, and the difficulty index. Thirteen of the 15 items were valid ($r \geq 0.297$), and the reliability coefficient was 0.71 (high). Among the valid items, nine showed good discrimination, one excellent, and three satisfactory; six items were of moderate difficulty and seven were difficult, with no easy items. Teachers can use the

13 items to assess students' diagnostic reasoning on electronic ignition faults, while the two rejected items need rewriting before use.

Keywords: assessment instrument, higher order thinking skills, electronic ignition system, 4D model, vocational high school

PENDAHULUAN

Kurikulum Sekolah Menengah Kejuruan (SMK) Program Keahlian Teknik dan Bisnis Sepeda Motor memuat mata pelajaran pemeliharaan kelistrikan sepeda motor. Salah satu kompetensi dasarnya menuntut siswa menganalisis sistem pengapian. Di bengkel, kompetensi ini berarti seorang teknisi mampu berangkat dari gejala, misalnya mesin sulit dihidupkan atau percikan api busi lemah, lalu menelusuri komponen yang menjadi sumber masalah: *pulser*, unit *Capacitor Discharge Ignition* (CDI), koil pengapian, kumparan pembangkit, atau busi. Pada sepeda motor yang beredar sekarang, sistem pengapian elektronik telah banyak menggantikan sistem pengapian konvensional, sehingga kemampuan mendiagnosis rangkaian elektronik ini menjadi bagian inti dari pekerjaan teknisi. Pekerjaan diagnosis semacam itu melibatkan dua proses kognitif, yaitu menganalisis dan mengevaluasi, yang dalam taksonomi Bloom revisi Anderson dan Krathwohl (2001) berada pada level C4 dan C5.

Keterampilan berpikir tingkat tinggi atau *Higher Order Thinking Skills* (HOTS) dipandang sebagai salah satu indikator keberhasilan pembelajaran abad ke-21 (Tan & Halili, 2015; Rahielanamy & Kamaludin, 2022). Capaian siswa Indonesia pada tugas yang menuntut penalaran masih rendah: pada *Programme for International Student Assessment* (PISA) 2018, Indonesia berada di peringkat 72 dari 77 negara pada bidang membaca (Schleicher, 2019). Di lingkungan pendidikan kejuruan,

Hasan dan Pardjono (2019) meneliti hubungan antara keterampilan berpikir tingkat tinggi dan kesiapan kerja siswa SMK, sedangkan Sukatiman dkk. (2020) mengembangkan pembelajaran berbasis masalah dengan *scaffolding* untuk meningkatkan keterampilan tersebut pada siswa vokasi. Upaya pembelajaran seperti ini hanya dapat dinilai keberhasilannya bila guru memiliki alat ukur yang benar-benar menguji level menganalisis dan mengevaluasi, bukan sekadar mengingat.

Observasi pendahuluan yang kami lakukan di sekolah menunjukkan keadaan sebaliknya. Soal pada materi sistem pengapian cenderung meminta siswa menyebutkan nama komponen, fungsi, atau urutan kerja, dan jarang menyajikan kasus kerusakan yang harus ditelusuri penyebabnya. Guru juga belum terbiasa menyusun butir soal yang menuntut penalaran tingkat tinggi. Akibatnya, nilai ulangan siswa tidak memberi informasi apakah mereka mampu mendiagnosis gangguan pengapian, padahal kemampuan itulah yang dituntut kriteria unjuk kerja di dunia kerja.

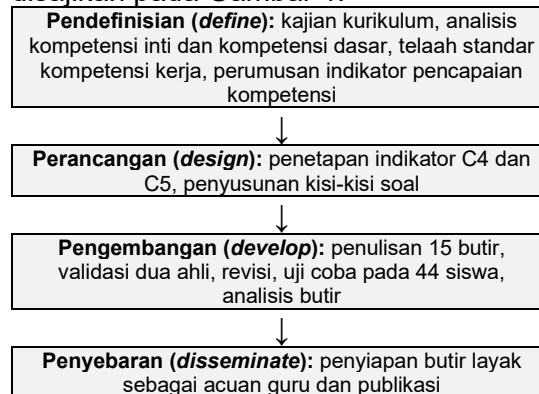
Penelitian pengembangan instrumen HOTS di Indonesia sudah cukup banyak, tetapi sebagian besar berada pada mata pelajaran umum. Pada matematika sekolah menengah pertama, Budiman dan Jailani (2014) menghasilkan 24 butir pilihan ganda dengan koefisien reliabilitas 0,713, Yunita dkk. (2018) menyusun butir level C4 sampai C6 yang diujikan di tiga sekolah di Jakarta, dan Nursalam dkk. (2018) memvalidasi isi instrumen dengan indeks validitas isi. Pada fisika, Sari dkk. (2018) memakai model 4D untuk tes pilihan ganda

berbasis pembelajaran elektronik, Najihah dkk. (2018) memakai model ADDIE (*Analyze, Design, Develop, Implement, Evaluate*) pada materi suhu dan kalor, Ramadhan dkk. (2019) menggabungkan teori tes klasik dengan model Rasch, Istiyono dkk. (2020) mengembangkan tes adaptif berbasis komputer, dan Irfansyah dan Sabani (2021) menyusun tes objektif hukum Newton berbasis pengetahuan konseptual. Kajian serupa juga dilakukan pada kimia (Rahielanamy & Kamaludin, 2022) dan ekonomi (Kurniawan & Lestari, 2019). Di SMK, pengembangan instrumen HOTS yang kami temukan masih terbatas pada rumpun bisnis dan perkantoran (Rahayu dkk., 2020; Rizal & Wulandari, 2020).

Dalam penelusuran kami, belum ditemukan instrumen HOTS yang diuji secara empiris untuk kompetensi diagnosis sistem pengapian elektronik sepeda motor. Konteks ini berbeda dari mata pelajaran umum. Butir soal harus memuat kasus kerusakan nyata, hasil pengukuran komponen, dan hubungan sebab-akibat dalam rangkaian listrik, sehingga temuan tentang kualitas butir pada matematika atau ekonomi belum tentu berlaku. Penelitian ini bertujuan menjelaskan tahapan pengembangan instrumen penilaian HOTS pada materi sistem pengapian elektronik sepeda motor dengan model 4D, serta menguji kelayakan instrumen tersebut berdasarkan validitas butir, reliabilitas, daya pembeda, dan tingkat kesukaran. Hasilnya berupa seperangkat butir yang dapat langsung dipakai guru Teknik dan Bisnis Sepeda Motor, disertai catatan tentang butir yang gagal dan alasan kegagalannya.

METODE

Penelitian ini merupakan penelitian dan pengembangan (*Research and Development*), yaitu penelitian yang menghasilkan produk tertentu sekaligus menguji keefektifan produk tersebut (Sugiyono, 2016). Pendekatan ini di bidang pembelajaran relatif baru mendapat perhatian dibandingkan pemakaiannya di sektor industri (Yuberti, 2014). Model yang dipakai adalah model 4D dari Thiagarajan dkk. (1974), yang terdiri atas tahap pendefinisian (*define*), perancangan (*design*), pengembangan (*develop*), dan penyebaran (*disseminate*). Kami memilih model ini karena tahapannya ringkas dan sudah dipakai untuk mengembangkan tes pilihan ganda HOTS (Sari dkk., 2018). Alur tahapan disajikan pada Gambar 1.



Gambar 1. Tahapan pengembangan instrumen dengan model 4D

Sumber: diadaptasi dari Thiagarajan dkk. (1974)

Penelitian berlangsung pada Februari–Agustus 2023 di dua SMK negeri di Jakarta yang membuka Program Keahlian Teknik dan Bisnis Sepeda Motor. Untuk menjaga kerahasiaan, kedua sekolah selanjutnya disebut Sekolah A dan Sekolah B. Subjek uji coba lapangan adalah 44 siswa kelas XII [PERLU DIPASTIKAN: jumlah siswa dari tiap sekolah dan teknik pengambilan sampel]. Siswa Sekolah A mengerjakan

tes secara daring melalui *Google Form* karena seluruh siswa kelas XII sedang menjalani praktik kerja lapangan, sedangkan siswa Sekolah B mengerjakan tes secara tatap muka dengan lembar soal cetak.

Pada tahap pendefinisian, kami mengkaji kurikulum yang berlaku di sekolah, menganalisis kompetensi inti dan kompetensi dasar untuk menemukan kompetensi yang menuntut berpikir tingkat tinggi, dan menelaah Standar Kompetensi Kerja Nasional Indonesia (SKKNI) subsektor sepeda motor pada bagian kriteria unjuk kerja untuk menetapkan batas materi soal. Hasil ketiga kajian itu dirumuskan menjadi indikator pencapaian kompetensi.

Pada tahap perancangan, kami menetapkan dua indikator berpikir tingkat tinggi dari taksonomi Bloom revisi, yaitu menganalisis (C4) dan mengevaluasi (C5) (Anderson & Krathwohl, 2001), lalu menyusun kisi-kisi soal. Kisi-kisi memetakan lima komponen sistem pengapian elektronik ke dalam 15 butir soal, seperti tercantum pada Tabel 1.

Tabel 1. Kisi-kisi instrumen penilaian berpikir tingkat tinggi

Komponen	Indikator soal	Nomor butir
<i>Pulser (pick-up coil)</i>	Disajikan permasalahan pada <i>pulser</i> ; siswa mendiagnosis penyebab dan akibat kerusakan	1, 2, 3
Busi	Disajikan permasalahan pada busi; siswa mendiagnosis penyebab dan akibat kerusakan	4, 5
Unit CDI	Disajikan permasalahan pada unit CDI; siswa mendiagnosis penyebab dan akibat kerusakan	6, 7
Koil pengapian	Disajikan permasalahan pada koil pengapian; siswa mendiagnosis penyebab dan akibat kerusakan. Disajikan gambar pengukuran koil;	8–13

	siswa menganalisis hasil pengukuran	
Kumparan pembangkit	Disajikan permasalahan pada kumparan pembangkit; siswa mendiagnosis penyebab dan akibat kerusakan	14, 15

Jumlah **15 butir**

Keterangan: kompetensi dasar yang diukur adalah menganalisis gangguan sistem pengapian elektronik, dengan indikator pencapaian kompetensi mendiagnosis gangguan, menemukan penyebab kerusakan, dan menyimpulkan dampak kerusakan pada komponen sistem pengapian elektronik. CDI = *Capacitor Discharge Ignition*.

Pada tahap pengembangan, kami menulis 15 butir pilihan ganda dengan lima pilihan jawaban sesuai kisi-kisi. Setiap butir dibuka dengan kasus gangguan atau gambar hasil pengukuran, kemudian siswa diminta menentukan penyebab, akibat, atau keputusan perbaikan. Kisi-kisi dan butir soal divalidasi oleh dua ahli, yaitu seorang dosen pendidikan teknik mesin dan seorang guru Teknik dan Bisnis Sepeda Motor di salah satu sekolah lokasi penelitian. Para ahli menilai kesesuaian butir dengan kompetensi dasar, indikator pencapaian kompetensi, materi, dan bahasa. Setelah direvisi, instrumen diujicobakan kepada siswa dan hasilnya dianalisis per butir. Pada tahap penyebaran, butir yang memenuhi kriteria disiapkan sebagai acuan bagi guru dan dipublikasikan melalui artikel ini.

Jawaban siswa diberi skor dikotomis, yaitu 1 untuk jawaban benar dan 0 untuk jawaban salah, lalu dianalisis dengan *Microsoft Excel* berdasarkan teori tes klasik. Validitas butir dihitung dengan koefisien korelasi *point biserial* antara skor butir dan skor total pada persamaan (1).

$$r_{pbi} = ((M_p - M_t) / S_t) \times \sqrt{(p / q)} \quad (1)$$

M_p adalah rerata skor total siswa yang menjawab benar pada butir tersebut, M_t rerata skor total, S_t simpangan baku

skor total, p proporsi siswa yang menjawab benar, dan $q = 1 - p$. Butir dinyatakan valid bila r hitung $\geq r$ tabel. Dengan jumlah responden 44 orang dan taraf signifikansi $\alpha = 0,05$, nilai r tabel adalah 0,297 (Sugiyono, 2016).

Reliabilitas diuji dari satu kali pemberian tes dengan pendekatan konsistensi internal (Aithal & Aithal, 2020), menggunakan rumus Kuder-Richardson 20 (KR-20) pada persamaan (2).

$$r_{11} = (k / (k - 1)) \times ((S_t^2 - \sum pq) / S_t^2) \quad (2)$$

Lambang k menyatakan jumlah butir, p proporsi siswa yang menjawab benar pada tiap butir, $q = 1 - p$, dan S_t^2 varians skor total. Koefisien ditafsirkan dengan klasifikasi 0,80–1,00 sangat tinggi, 0,60–0,79 tinggi, 0,40–0,59 cukup, 0,20–0,39 rendah, dan 0,00–0,19 sangat rendah (Hamidah, 2015).

Daya pembeda dihitung setelah siswa dibagi menjadi kelompok atas dan kelompok bawah berdasarkan skor total [PERLU DIPASTIKAN: proporsi pembagian kelompok], menggunakan persamaan (3).

$$D = (JB_a / JS_a) - (JB_b / JS_b) \quad (3)$$

JB_a dan JB_b adalah jumlah siswa kelompok atas dan kelompok bawah yang menjawab benar, sedangkan JS_a dan JS_b adalah jumlah siswa pada masing-masing kelompok. Indeks 0,00–0,20 tergolong jelek, 0,21–0,40 cukup, 0,41–0,70 baik, 0,71–1,00 baik sekali, dan nilai negatif tidak baik (Arikunto, 2017). Tingkat kesukaran dihitung dengan persamaan (4).

$$P = B / JS \quad (4)$$

B adalah jumlah siswa yang menjawab benar dan JS jumlah seluruh siswa peserta tes. Indeks 0,00–0,30 tergolong sukar, 0,31–0,70 sedang, dan 0,71–1,00 mudah (Arikunto, 2017).

HASIL DAN PEMBAHASAN

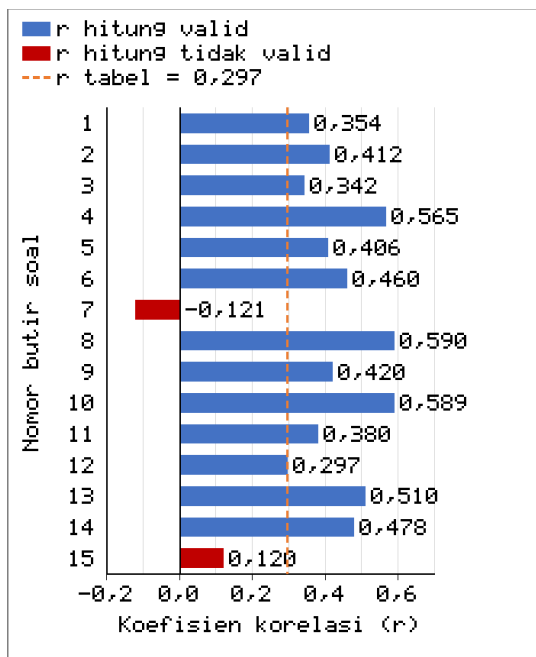
Hasil pengembangan instrumen

Tahap pendefinisian menghasilkan satu kompetensi dasar yang menuntut berpikir tingkat tinggi, yaitu menganalisis gangguan sistem pengapian elektronik. Kompetensi ini diturunkan menjadi tiga indikator pencapaian kompetensi: mendiagnosis gangguan, menemukan penyebab kerusakan, dan menyimpulkan dampak kerusakan pada komponen. Tahap perancangan menghasilkan kisi-kisi pada Tabel 1. Porsi terbesar jatuh pada koil pengapian (enam butir) karena komponen ini diuji melalui dua jenis tugas, yaitu diagnosis kasus kerusakan dan pembacaan hasil pengukuran.

Kedua validator menyatakan instrumen layak digunakan dengan sedikit perbaikan, yaitu mengganti rumusan capaian pembelajaran pada kisi-kisi menjadi indikator pencapaian kompetensi. Setelah perbaikan tersebut, instrumen langsung diujicobakan kepada siswa. Penilaian ahli dalam penelitian ini bersifat kualitatif dan tidak menghasilkan indeks angka.

Validitas dan reliabilitas butir

Grafik 1 menyajikan koefisien korelasi point biserial setiap butir terhadap batas r tabel 0,297. Tiga belas butir memenuhi kriteria valid dengan r hitung antara 0,297 dan 0,590. Korelasi tertinggi terdapat pada butir 8 (0,590) dan butir 10 (0,589), keduanya tentang koil pengapian, sedangkan butir 12 tepat berada pada batas (0,297). Dua butir tidak valid, yaitu butir 7 tentang unit CDI yang berkorelasi negatif (–0,121) dan butir 15 tentang kumparan pembangkit (0,120).



Grafik 1. Koefisien korelasi point biserial tiap butir soal (N = 44)

Keterangan: garis putus-putus menandai r tabel 0,297 ($\alpha = 0,05$); batang merah menandai butir tidak valid.

Perhitungan KR-20 menghasilkan koefisien reliabilitas 0,71, yang termasuk kategori tinggi menurut klasifikasi Hamidah (2015) [PERLU DIPASTIKAN: apakah koefisien dihitung dari 15 butir atau dari 13 butir valid].

Daya pembeda dan tingkat kesukaran

Analisis daya pembeda dan tingkat kesukaran dilakukan pada 13 butir yang valid (Tabel 2). Tidak ada butir dengan daya pembeda jelek atau negatif. Sembilan butir berdaya pembeda baik, satu butir (butir 8) baik sekali, dan tiga butir (butir 1, 3, dan 12) cukup. Dari sisi tingkat kesukaran, tujuh butir tergolong sukar, enam butir sedang, dan tidak ada butir mudah. Dua dari tiga butir berdaya pembeda cukup, yaitu butir 3 dan 12, juga tergolong sukar.

Tabel 2. Hasil analisis butir soal

Butir	Komponen	Daya pembeda	Tingkat kesukaran
1	Pulser	Cukup	Sedang
2	Pulser	Baik	Sedang
3	Pulser	Cukup	Sukar
4	Busi	Baik	Sedang
5	Busi	Baik	Sukar
6	Unit CDI	Baik	Sukar
7	Unit CDI	Tidak valid	–
8	Koil pengapian	Baik sekali	Sedang
9	Koil pengapian	Baik	Sukar
10	Koil pengapian	Baik	Sedang
11	Koil pengapian	Baik	Sukar
12	Koil pengapian	Cukup	Sukar
13	Koil pengapian	Baik	Sedang
14	Kumparan pembangkit	Baik	Sukar
15	Kumparan pembangkit	Tidak valid	–

Keterangan: butir 7 dan 15 tidak valid sehingga tidak dianalisis lebih lanjut (–). CDI = *Capacitor Discharge Ignition*.

Pembahasan

Model 4D menghasilkan instrumen yang sebagian besar butirnya layak dipakai. Tiga belas dari lima belas butir valid, dan koefisien reliabilitasnya berada pada kategori tinggi. Nilai reliabilitas ini setara dengan temuan Budiman dan Jailani (2014) untuk soal pilihan ganda HOTS matematika (0,713), sedikit di bawah Kurniawan dan Lestari (2019) yang memperoleh 0,786 dan 0,763 dari 15 butir HOTS ekonomi, dan di bawah Rizal dan Wulandari (2020) yang melaporkan 0,820 untuk 20 butir pilihan ganda dan 5 butir uraian di SMK. Reliabilitas konsistensi internal cenderung naik bila jumlah butir bertambah, sehingga nilai 0,71 dari instrumen sependek ini masih dapat diterima. Bila skor akan dipakai untuk keputusan kelulusan, guru sebaiknya menambah butir yang setara daripada mengandalkan 13 butir saja.

Sebaran tingkat kesukaran instrumen ini lebih berat dibandingkan penelitian sejenis. Rahielanamy dan Kamaludin (2022) memperoleh 15% butir mudah, 75% sedang, dan 10% sukar,

sedangkan Budiman dan Jailani (2014) melaporkan rerata tingkat kesukaran 0,406 atau kategori sedang. Pada penelitian ini, 7 dari 13 butir valid tergolong sukar. Kami menduga ada tiga penyebab. Pertama, setiap butir menuntut siswa menghubungkan gejala dengan komponen dalam rangkaian, misalnya menafsirkan hasil pengukuran koil, dan instrumen tidak memuat butir level menerapkan yang biasanya berfungsi sebagai butir pembuka yang lebih ringan. Kedua, siswa Sekolah A mengerjakan tes secara daring ketika sedang praktik kerja lapangan, sehingga kondisi pengerjaannya tidak setara dengan siswa yang diuji tatap muka. Ketiga, siswa terbiasa dengan soal hafalan, sesuai temuan observasi pendahuluan. Data penelitian ini belum cukup untuk memisahkan pengaruh ketiga faktor tersebut karena hasil tidak dianalisis menurut sekolah atau moda tes.

Daya pembeda butir menunjukkan bahwa instrumen mampu memisahkan siswa yang sudah menguasai diagnosis pengapian dari siswa yang belum. Rahielanamy dan Kamaludin (2022) mendapatkan 85% butir berdaya pembeda baik, sedangkan pada penelitian ini 10 dari 13 butir valid (77%) berdaya pembeda baik atau baik sekali. Butir 3 dan 12 perlu dicermati. Keduanya tergolong sukar dan hanya berdaya pembeda cukup, yang berarti siswa kelompok atas pun banyak yang menjawab salah. Butir 12 juga memiliki r hitung yang sama persis dengan r tabel. Stimulus dan kunci jawaban kedua butir ini sebaiknya ditelaah ulang sebelum dimasukkan ke bank soal.

Korelasi negatif pada butir 7 menunjukkan bahwa siswa dengan skor total tinggi justru lebih sering salah

menjawab butir tersebut. Pola seperti ini biasanya muncul bila kunci jawaban meragukan, ada pengecoh yang tampak lebih benar daripada kunci bagi siswa yang menguasai materi, atau stimulus dapat ditafsirkan ganda. Gugurnya butir 7 dan 15 juga mengurangi keterwakilan materi: unit CDI dan kumparan pembangkit kini masing-masing hanya diwakili satu butir, sementara koil pengapian diwakili enam butir. Penelitian ini belum menganalisis keberfungsian pengecoh, padahal analisis tersebut dipakai Sari dkk. (2018) dan Rizal dan Wulandari (2020) untuk menilai kualitas pilihan jawaban. Untuk butir 7 dan 15, analisis pengecoh menjadi langkah pertama sebelum butir direvisi atau ditulis ulang.

Beberapa keterbatasan membatasi tafsir hasil ini. Validasi ahli dilakukan secara kualitatif oleh dua orang, sedangkan penelitian lain mengukur validitas isi dengan indeks angka, misalnya *Content Validity Index* dan *Content Validity Ratio* (Nursalam dkk., 2018), persentase kelayakan dari validator (Rahayu dkk., 2020), atau formula Aiken (Ramadhan dkk., 2019). Analisis butir juga memakai teori tes klasik dengan 44 responden, sehingga indeks daya pembeda dan tingkat kesukaran bergantung pada kelompok siswa yang diuji. Ramadhan dkk. (2019) menggabungkan teori tes klasik dengan model Rasch untuk memperoleh bukti empiris kualitas butir, dan Istiyono dkk. (2020) mengembangkan butir HOTS menjadi tes adaptif berbasis komputer. Selain itu, instrumen ini baru mencakup level C4 dan C5 dan belum menyentuh level mencipta (C6).

Bagi guru Teknik dan Bisnis Sepeda Motor, ketiga belas butir valid dapat

dipakai sebagai tes diagnostik setelah pembelajaran sistem pengapian elektronik. Pola jawaban siswa menunjukkan komponen yang belum dikuasai, misalnya bila banyak siswa salah pada butir koil pengapian yang menyajikan hasil pengukuran. Instrumen ini juga dapat dipasangkan dengan model pembelajaran yang dirancang untuk melatih berpikir tingkat tinggi, seperti pembelajaran berbasis masalah dengan *scaffolding* (Sukatiman dkk., 2020), sehingga guru dapat menilai dampak pembelajaran tersebut dengan alat ukur yang sesuai konteks otomotif. Hasan dan Pardjono (2019) yang meneliti hubungan berpikir tingkat tinggi dengan kesiapan kerja siswa SMK memberi alasan tambahan agar penilaian di sekolah kejuruan tidak berhenti pada soal hafalan.

KESIMPULAN

Model pengembangan Thiagarajan yang terdiri atas tahap pendefinisian, perancangan, pengembangan, dan penyebaran dapat dipakai untuk menyusun instrumen penilaian berpikir tingkat tinggi pada materi sistem pengapian elektronik sepeda motor. Tahapan yang dilalui meliputi kajian kurikulum dan kriteria unjuk kerja, perumusan indikator pencapaian kompetensi, penyusunan kisi-kisi pada level menganalisis dan mengevaluasi, penulisan butir berbasis kasus kerusakan, validasi ahli, uji coba lapangan, dan analisis butir. Sebagian besar butir yang dihasilkan valid, instrumen memiliki reliabilitas tinggi, butir yang valid mampu membedakan siswa kelompok atas dan kelompok bawah, dan sebaran tingkat kesukarannya condong ke kategori sukar. Instrumen ini layak dipakai guru untuk mengukur kemampuan siswa

mendiagnosis gangguan pengapian elektronik. Butir yang tidak valid perlu ditulis ulang setelah analisis pengecoh, dan penelitian lanjutan perlu menambah butir untuk unit CDI dan kumparan pembangkit, memakai indeks validitas isi yang terukur, melibatkan lebih banyak sekolah, serta menguji parameter butir dengan model Rasch.

DAFTAR PUSTAKA

- Aithal, A., & Aithal, P. S. (2020). Development and validation of survey questionnaire & experimental data – A systematical review-based statistical approach. *International Journal of Management, Technology, and Social Sciences*, 5(2), 233–251. <https://doi.org/10.47992/IJMTS.2581.6012.0116>
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Addison Wesley Longman.
- Arikunto, S. (2017). *Dasar-dasar evaluasi pendidikan* (Edisi 2). Bumi Aksara.
- Budiman, A., & Jailani, J. (2014). Pengembangan instrumen asesmen higher order thinking skill (HOTS) pada mata pelajaran matematika SMP kelas VIII semester 1. *Jurnal Riset Pendidikan Matematika*, 1(2), 139–151. <https://doi.org/10.21831/jrpm.v1i2.2671>
- Hamidah. (2015). Pengembangan instrumen penilaian. *Prodi P.ADP FF UNY*, 13(1), 1–100.
- Hasan, A., & Pardjono, P. (2019). The correlation of higher order thinking skills and work readiness of vocational high school students.

- Jurnal Pendidikan Teknologi dan Kejuruan*, 25(1), 52–61.
<https://doi.org/10.21831/jptk.v25i1.19118>
- Irfansyah, I., & Sabani, S. (2021). Pengembangan tes objektif fisika SMA hukum Newton berbasis pengetahuan konseptual. *Journal of Natural Sciences*, 2(3), 95–104.
<https://doi.org/10.34007/jonas.v2i3.149>
- Istiyono, E., Dwandaru, W. S. B., Setiawan, R., & Megawati, I. (2020). Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use. *European Journal of Educational Research*, 9(1), 91–101. <https://doi.org/10.12973/euler.9.1.91>
- Kurniawan, R. Y., & Lestari, D. (2019). The development assessment instruments of higher order thinking skills on economic subject. *Dinamika Pendidikan*, 14(1), 102–115.
<https://doi.org/10.15294/dp.v14i1.19226>
- Najihah, A. R., Serevina, V., & Delina, M. (2018). The development of high order thinking skills (HOTS) assessment instrument for temperature and heat learning. *JPPPF: Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 4(1), 19–26.
<https://doi.org/10.21009/1.04103>
- Nursalam, Angriani, A. D., Darmawati, Baharuddin, & Aminuddin. (2018). Developing test instruments for measurement of students' high-order thinking skill on mathematics in junior high school in Makassar. *Journal of Physics: Conference Series*, 1028, 012169.
<https://doi.org/10.1088/1742-6596/1028/1/012169>
- Rahayu, W. P., Hasanah, U., & Wiliandri, R. (2020). Developing a higher order thinking skill-oriented and metacognitive-based assessment for vocational school students. *Jurnal Pendidikan Bisnis dan Manajemen*, 6(1), 10–23.
<https://doi.org/10.17977/um003v6i12020p010>
- Rahielanamy, R., & Kamaludin, A. (2022). Development of higher order thinking skill qualified test on colligative properties of solutions. *Jurnal Pendidikan Sains Indonesia*, 10(2), 294–311.
<https://doi.org/10.24815/jpsi.v10i2.23968>
- Ramadhan, S., Mardapi, D., Prasetyo, Z. K., & Utomo, H. B. (2019). The development of an instrument to measure the higher order thinking skill in physics. *European Journal of Educational Research*, 8(3), 743–751. <https://doi.org/10.12973/euler.8.3.743>
- Rizal, A. N., & Wulandari, S. S. (2020). Pengembangan instrumen penilaian higher order thinking skills (HOTS) mata pelajaran otomatisasi tata kelola humas dan keprotokolan di SMK Negeri Mojoagung. *Jurnal Pendidikan Administrasi Perkantoran*, 8(2), 194–204.
<https://doi.org/10.26740/jpap.v8n2.p194-204>
- Sari, D. R. U., Wahyuni, S., & Bachtiar, R. W. (2018). Pengembangan instrumen tes multiple choice high order thinking pada pembelajaran fisika berbasis e-learning di SMA.

- Jurnal Pembelajaran Fisika*, 7(1), 100.
<https://doi.org/10.19184/jpf.v7i1.7231>
- Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. OECD.
- Sugiyono. (2016). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Sukatiman, S., Akhyar, M., Siswandari, S., & Roemintoyo, R. (2020). Enhancing higher-order thinking skills in vocational education through scaffolding-problem based learning. *Open Engineering*, 10(1), 612–619.
<https://doi.org/10.1515/eng-2020-0070>
- Tan, S. Y., & Halili, S. H. (2015). Effective teaching of higher-order thinking (HOT) in education. *The Online Journal of Distance Education and E-Learning*, 3(2), 41–47.
- Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Leadership Training Institute/Special Education, University of Minnesota.
- Yuberti. (2014). "Penelitian dan pengembangan" yang belum diminati dan perspektifnya. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, 3(2), 1–15.
<https://doi.org/10.24042/jpifalbiruni.v3i2.69>
- Yunita, Y., Wahidin, W., & Tsurayya, A. (2018). The development of mathematics higher order thinking skills instrument for grade VIII junior high school. *Journal of Physics: Conference Series*, 1088, 012093.
<https://doi.org/10.1088/1742-6596/1088/1/012093>