ASIAN PENDIDIKAN VOL. 6 NO. 2 (2026) 88–105

© Association of Researcher of Skills and Vocational Training, Malaysia

ASIAN PENDIDIKAN
ISSN: 2735-2331
DOI: https://doi.org/10.53797/aspen.v6i2.11.2026

Keperluan Pembangunan Instrumen Penilaian Pengajaran Pelatih Internship Program TVET-I Berasaskan Metodologi Pengajaran 4-Langkah: Analisis Dokumentasi

(Needs for Developing a Teaching Assessment Instrument for TVET-I Internship Trainees Based on the 4-Step Teaching Methodology: A Documentary Analysis)

Zulkifli Mohd Sidi[1,*], Bizaisham Bidin[2], Mohd Hazimin Mohd Nawawi[3] & Iskandar Kamarudin[4]

[1]Persatuan Penyelidik Latihan Kemahiran & Vokasional (ARSVOT), Shah Alam, Selangor, MALAYSIA

[2]Jabatan Pembangunan Kemahiran (JPK), Putrajaya, Wilayah Persekutuan, MALAYSIA

[3]Kampus Satelit CIAST (CSC), ADTEC JTM Kampus Marang, Terengganu, MALAYSIA

[4]Pusat Latihan Pengajar dan Kemahiran Lanjutan (CIAST), Shah Alam, Selangor, MALAYSIA

*Corresponding Author: zulkiflisidi@gmail.com

Received 22 October 2026; Accepted 15 November2026; Available online 21 November2026

Abstract: The transition of skills instructor qualifications in Malaysia from the National Occupational Skills Standard (NOSS) for Vocational Training Operation (VTO) Level 3 (I-031-3:2014, Malaysian Skills Certificate, SKM) to TVET Instruction (TVET-I) Level 4 (P853-002-4:2021, Malaysian Skills Diploma, DKM) has raised the demand for competency-based pedagogical mastery, including a mandatory internship at an Accredited Centre (PB). This first-phase study uses documentary analysis to build the justification for a follow-up study that will develop and validate an assessment instrument based on the 4-Step Teaching Methodology. A cross-document analysis was conducted on current standards and regulations, teaching-material development guidelines, internship visit guidelines, official assessment forms, placement records and related assessment literature. The teaching simulation forms used by Internship Supervisors, Assessment Officers (PP) and External Verification Officers (PPL) show five core measurement weaknesses: (1) a generic 0–3 rating scale without behavioural descriptors; (2) no exclusion codes ('Insufficient Evidence' and 'Not Applicable') in the 30-minute Teaching Simulation; (3) a mixed construct, as the quality of the Written Instructional Material (WIM) is scored again although it is assessed separately; (4) identical 43 items and weightings for theory (time ratio 3:2:1) and practical (time ratio 1:3:2) sessions; and (5) inconsistent final weightings across assessors. These findings provide an evidence-based justification for developing an analytic rubric, a scoring manual and rater calibration, which must then be tested empirically before the official form is revised.

Keywords: Internship; Accredited Centre; TVET-I teaching assessment; 4-Step Teaching Methodology; micro-teaching

Abstrak: Peralihan kelayakan pengajar kemahiran di Malaysia daripada Standard Kemahiran Pekerjaan Kebangsaan (NOSS) *Vocational Training Operation* (VTO) Tahap 3 (I-031-3:2014, Sijil Kemahiran Malaysia, SKM) kepada Pengajaran TVET (TVET-I) Tahap 4 (P853-002-4:2021, Diploma Kemahiran Malaysia, DKM) meningkatkan tuntutan penguasaan pedagogi berasaskan kompetensi, termasuk kewajipan latihan internship di Pusat Bertauliah (PB). Kajian fasa pertama ini menggunakan pendekatan analisis dokumentasi untuk membina justifikasi bagi kajian lanjutan yang akan membangunkan dan mengesahkan instrumen penilaian berasaskan Metodologi Pengajaran 4-Langkah. Analisis silang dilaksanakan terhadap standard dan peraturan yang berkuat kuasa, panduan pembangunan bahan pengajaran, panduan operasi lawatan internship, borang penilaian rasmi, rekod penempatan serta literatur penilaian yang berkaitan. Borang Simulasi Mengajar yang digunakan oleh Penyelia Internship, Pegawai Penilai (PP) dan Pegawai Pengesah Luaran (PPL) didapati mempunyai lima kelemahan pengukuran teras: (1) skala umum 0–3 tanpa deskriptor tingkah laku; (2) ketiadaan kod pengecualian 'Bukti Tidak Mencukupi' dan 'Tidak Berkenaan'

Corresponding author: zulkiflisidi@gmail.com

dalam Simulasi Mengajar 30 minit; (3) pencampuran konstruk apabila kualiti *Written Instructional Material* (WIM) dinilai semula walaupun telah dinilai secara berasingan; (4) penggunaan 43 item dan wajaran yang sama bagi sesi teori (nisbah masa 3:2:1) dan amali (nisbah masa 1:3:2); dan (5) pemberat akhir yang tidak selaras antara penilai. Dapatan ini menyediakan justifikasi berasaskan bukti bagi pembangunan rubrik analitik, manual penskoran dan penentukuran penilai, yang perlu diuji secara empirikal sebelum borang rasmi disemak semula.

Kata Kunci: Internship; Pusat Bertauliah; penilaian pengajaran TVET-I; Metodologi Pengajaran 4-Langkah; Simulasi Mengajar

1. Pengenalan

1.1. Latar Belakang

Keberkesanan sistem Pendidikan dan Latihan Teknikal dan Vokasional (TVET) bergantung pada keupayaan tenaga pengajar mengaplikasikan pengetahuan dan kemahiran teknikal melalui pengajaran yang sistematik, selamat dan berpusatkan pelatih (*trainee-centred*), sama ada di bilik kuliah mahupun di bengkel. Dalam Sistem Persijilan Kemahiran Malaysia (SPKM) di bawah Jabatan Pembangunan Kemahiran (JPK), pembangunan kompetensi pengajar telah diperkukuh melalui pengenalan program Pengajaran TVET (TVET-I) berasaskan NOSS P853-002-4:2021 pada peringkat Diploma Kemahiran Malaysia (DKM) Tahap 4. Program ini menggantikan program terdahulu, iaitu *Vocational Training Operation* (VTO), NOSS I-031-3:2014, pada peringkat Sijil Kemahiran Malaysia (SKM) Tahap 3 (Jabatan Pembangunan Kemahiran [JPK], 2021).

Penstrukturan semula ini bertujuan memperkukuh kompetensi pengajar agar selaras dengan keperluan pengajaran dan latihan kemahiran semasa. Melalui program TVET-I, calon pengajar perlu melengkapkan enam Unit Kompetensi teras (C01 hingga C06) serta menjalani latihan internship di Pusat Bertauliah (PB). Latihan internship memberikan peluang kepada calon pengajar untuk mengaplikasikan pengetahuan, kemahiran pengajaran dan amalan profesional dalam situasi latihan sebenar. Pelaksanaan dan penilaian internship juga menjadi komponen penting dalam menentukan pencapaian kompetensi mengajar sebelum calon memenuhi syarat persijilan sebagai pengajar TVET.

Isu kualiti pengajar vokasional turut menjadi perhatian antarabangsa. OECD (2021) menegaskan bahawa pengajar TVET perlu menggabungkan kepakaran industri dengan kemahiran pedagogi, manakala Smith dan Tuck (2023) mendapati kelayakan pedagogi pengajar vokasional mempengaruhi amalan pengajaran mereka. Tinjauan konseptual Antera (2021) pula menunjukkan bahawa kompetensi profesional pengajar vokasional jarang didefinisikan secara jelas dalam kajian, sedangkan definisi konstruk yang jelas merupakan prasyarat kepada sebarang instrumen penilaian. Di Malaysia, Ismail et al. (2017) mendapati sijil VTO terbitan CIAST/JPK menjadi sumber utama kompetensi pengajar TVET, manakala tinjauan skop Saipudin dan Suhairom (2021) mengenal pasti ketidakseimbangan antara kompetensi teknikal dengan kompetensi bukan teknikal serta pendekatan pedagogi yang lemah dalam kalangan pendidik TVET. Dapatan ini menunjukkan bahawa penilaian kompetensi pedagogi pengajar TVET wajar diberi perhatian khusus.

Dalam artikel ini, ‘pelatih internship’ atau ‘pelatih TVET-I’ merujuk calon pengajar yang dinilai semasa internship, manakala ‘pelatih’ dalam huraian item Simulasi Mengajar merujuk peserta latihan yang diajar oleh calon pengajar tersebut.

1.2. Justifikasi & Pendekatan Kajian

Kajian ini mengambil kira dua pertimbangan utama bagi memperkukuh asas pembangunan instrumen penilaian internship TVET-I. Pertama, borang penilaian rasmi bagi program P853-002-4:2021 telah pun digunakan dalam penilaian internship dan mencerminkan amalan penilaian sedia ada dalam SPKM. Dokumentasi dan pengalaman pelaksanaannya menyediakan sumber maklumat yang bernilai untuk diteliti secara sistematik bagi mengenal pasti kekuatan, kesesuaian dan ruang penambahbaikan instrumen. Kedua, borang rasmi tersebut disokong oleh manual dan prosedur pelaksanaan yang menyediakan landasan pentadbiran berstruktur, sekali gus membuka peluang untuk memperkukuh aspek kesahan, kebolehpercayaan, ketekalan dan keupayaan diagnostik instrumen dalam menilai kompetensi pedagogi pelatih TVET-I.

Sehubungan itu, pendekatan Penyelidikan Reka Bentuk dan Pembangunan (*Design and Development Research*, DDR) bagi pembangunan produk dan alat (Jenis 1), seperti yang dihuraikan oleh Richey dan Klein (2007), merupakan pendekatan metodologi yang sesuai. Pendekatan ini membolehkan keperluan pembangunan dan penambahbaikan instrumen dikenal pasti melalui analisis sistematik terhadap dokumen piawaian, instrumen penilaian dan prosedur pelaksanaan sedia ada, dan pada fasa seterusnya, maklumat daripada pihak berkepentingan. Hasil analisis tersebut menjadi asas kepada pembangunan instrumen penilaian internship TVET-I yang berlandaskan bukti dan memenuhi keperluan penilaian kompetensi pedagogi pelatih. Melalui pendekatan ini, instrumen yang dibangunkan dijangka lebih sistematik, komprehensif dan berupaya menghasilkan maklumat diagnostik yang bermakna tentang tahap pencapaian, kekuatan serta aspek kompetensi pengajaran yang perlu ditambah baik.

Penelitian awal terhadap dokumen pelaksanaan internship TVET-I menunjukkan bahawa prosedur penilaian telah ditetapkan secara berstruktur, merangkumi Simulasi Mengajar selama 30 minit, markah lulus minimum 60% dan penilaian semula (ulang nilai) dalam tempoh dua minggu bagi pelatih yang belum mencapai tahap kelulusan. Tempoh

simulasi ini boleh menjadi asas pemerhatian kompetensi pengajaran, namun Hill et al. (2012) mengingatkan bahawa skor pemerhatian pengajaran dipengaruhi bukan sahaja oleh penilai, malah oleh bilangan dan tempoh sesi yang diperhatikan. Walaupun prosedur tersebut menyediakan panduan kepada Penyelia Internship (PI), Pegawai Penilai (PP) CIAST/CSC dan Pegawai Pengesah Luaran (PPL), beberapa aspek instrumen masih memerlukan penelitian. Antaranya ialah kejelasan deskriptor tingkah laku bagi skala 0–3, kaedah merekod item yang tidak dapat diperhatikan, pengasingan penilaian dokumen *Written Instructional Material* (WIM) daripada prestasi pengajaran, serta kesesuaian item dan wajaran penilaian antara sesi teori (nisbah masa 3:2:1) dengan amali (1:3:2). Selain itu, struktur dalaman instrumen dan ketekalan pemarkahan antara penilai perlu diuji secara empirikal bagi menentukan kesahan dan kebolehpercayaan keputusan penilaian.

Aspek-aspek tersebut perlu diteliti secara sistematik melalui analisis dokumen dan, pada kajian seterusnya, pengumpulan bukti empirikal. Kajian tempatan tentang penilaian amalan mengajar, seperti analisis rekod borang penilaian latihan mengajar oleh Lee et al. (2020), menumpukan taburan skor calon guru dan bukannya kualiti reka bentuk instrumen itu sendiri. Di peringkat antarabangsa pula, Gitomer et al. (2021) menunjukkan bahawa penilaian prestasi mengajar berimpak tinggi bagi guru pelatih boleh mempunyai kelemahan reka bentuk pemarkahan dan kebolehpercayaan yang serius walaupun digunakan secara meluas. Setakat penelitian pengkaji, belum ada kajian yang menilai secara sistematik reka bentuk borang Simulasi Mengajar internship TVET-I terhadap prinsip rubrik analitik dan piawaian pengukuran. Jurang inilah yang ditangani oleh kajian ini. Kajian ini menggunakan pendekatan analisis dokumen bagi menyediakan justifikasi saintifik dan teknikal terhadap keperluan kajian lanjutan (selanjutnya disebut kajian utama) untuk membangunkan instrumen penilaian pengajaran pelatih internship TVET-I. Penelitian tertumpu kepada perubahan keperluan persijilan pengajar, struktur dan prosedur penilaian sedia ada, kesesuaian item pengukuran serta keperluan penambahbaikan instrumen.

Berdasarkan jurang tersebut, kajian ini menganalisis dokumentasi rasmi dan literatur berkaitan untuk mencapai empat tujuan, iaitu (i) menganalisis perubahan keperluan persijilan dan kompetensi pedagogi pengajar berikutan peralihan daripada program VTO Tahap 3 (SKM) kepada TVET-I Tahap 4 (DKM); (ii) meneliti kesesuaian struktur item, skala pemarkahan, padanan item dengan konstruk yang dinilai dan ketepatan teknikal pemarkahan borang penilaian rasmi; (iii) menilai kesesuaian pelaksanaan dan pengukuran bagi perbezaan wajaran penilaian sesi teori dan amali, tempoh simulasi, prosedur penilaian semula (ulang nilai) serta kepelbagaian konteks internship; dan (iv) membangunkan matriks pemetaan item penilaian sedia ada berdasarkan empat komponen utama Metodologi Pengajaran 4-Langkah. Tujuan pertama dijawab dalam Bahagian 4.1, tujuan kedua dalam Bahagian 4.2, tujuan ketiga dalam Bahagian 4.3 dan 4.5, manakala tujuan keempat dalam Bahagian 4.4. Dengan menggandingkan dokumen peraturan latihan kemahiran Malaysia bersama landasan teori pedagogi dan pengukuran pendidikan, bahagian seterusnya membentangkan kerangka konseptual yang memandu analisis ini.

2. Landasan Konseptual & Tinjauan Literatur Berfokus

Penilaian pengajaran dalam latihan kemahiran menuntut kesepaduan antara teori pedagogi vokasional dengan piawaian pengukuran pendidikan agar skor yang diperoleh mencerminkan kecekapan mengajar sebenar, dan bukannya kesan ralat instrumen atau tafsiran subjektif penilai. Bahagian ini mensintesis empat teras literatur yang membina hujah keperluan kepada kajian lanjutan yang lebih menyeluruh.

2.1. Pedagogi 4-Langkah, *Laws of Learning* & Pengetahuan Kandungan Pedagogi

Penguasaan kemahiran teknikal tidak semestinya menjamin keupayaan seseorang pengajar menyampaikan pengetahuan dan kemahiran tersebut secara berkesan kepada pelatih. Shulman (1986) memperkenalkan konsep Pengetahuan Kandungan Pedagogi (*Pedagogical Content Knowledge*, PCK) yang menekankan keupayaan pengajar menggabungkan pengetahuan bidang dengan kemahiran pedagogi bagi membantu pelatih memahami dan menguasai sesuatu pembelajaran; konsep ini terus digunakan secara meluas dalam pendidikan pasca menengah (Sarkar et al., 2024). Keupayaan ini merangkumi penyampaian penerangan yang tersusun, penggunaan contoh dan analogi yang sesuai, demonstrasi langkah kerja serta pemberian bimbingan dan pembetulan apabila pelatih menghadapi kesukaran memahami konsep atau melaksanakan prosedur kerja (OECD, 2021).

Dalam konteks latihan berasaskan kompetensi (*Competency-Based Training*, CBT) di Malaysia, Zulkifli (2019) dan Bidin et al. (2020) membincangkan pelaksanaan Metodologi Pengajaran 4-Langkah (*4-Step Teaching Pattern*) dalam latihan pengajar kemahiran, khususnya amalan pengajaran di Pusat Latihan Pengajar dan Kemahiran Lanjutan (CIAST). Kaedah ini mempunyai akar sejarah dalam pendekatan pengajaran berperingkat yang dikaitkan dengan Herbart (1904) dan Allen (1919). Pelaksanaannya juga boleh dihuraikan melalui prinsip Undang-undang Pembelajaran (*Laws of Learning*). Langkah Persediaan menekankan kesediaan pelatih untuk menerima pembelajaran (*Law of Readiness*), manakala langkah Penyampaian memberikan tumpuan kepada penerangan dan demonstrasi yang jelas, menarik serta berkesan, selaras dengan prinsip *Law of Effect*, *Law of Primacy* dan *Law of Intensity*. Seterusnya, langkah Penggunaan memberi peluang kepada pelatih mempraktikkan kemahiran melalui latihan, pemantauan, bimbingan dan pembetulan kesilapan secara berterusan, bertepatan dengan *Law of Exercise*. Akhir sekali, langkah Pengesahan bertujuan menentukan tahap penguasaan pengetahuan dan kemahiran pelatih melalui tugasan atau ujian secara individu tanpa bantuan pengajar, yang boleh dikaitkan dengan *Law of Effect* dan *Law of Exercise*.

Bidin et al. (2020) turut menghuraikan lima kriteria penilaian penyampaian pengajaran bagi program VTO Tahap 3 melalui borang JPK/I031/3D-1, iaitu *Preparation*, *Presentation*, *Application*, *Confirmation* dan *General Teaching Techniques*. Kriteria tersebut merangkumi aspek pelaksanaan pengajaran serta perkara berkaitan pembangunan WIM. Walau bagaimanapun, perbincangan dalam artikel tersebut lebih tertumpu kepada kerangka konseptual penilaian dan tidak memperincikan analisis empirikal berkaitan taburan skor, kebolehpercayaan antara penilai atau kesahan struktur faktor instrumen. Keadaan ini menunjukkan bahawa kualiti pengukuran masih memerlukan penelitian lanjut, khususnya apabila kriteria penilaian terdahulu dijadikan rujukan dalam pembangunan instrumen bagi program TVET-I.

Selain itu, penyelidikan Zulkifli (2018) berasaskan pendekatan DDR telah menghasilkan model pentaksiran kompetensi holistik bagi sektor TVET. Model tersebut menekankan tiga orientasi perkembangan kompetensi, iaitu pemupukan, pemantapan dan penghayatan, dan dibangunkan melalui Teknik Delphi Diubah Suai (*Modified Delphi Technique*, MDT) dan Kaedah Delphi Kabur (*Fuzzy Delphi Method*, FDM). Walaupun kajian tersebut menyumbang kepada pembangunan model pentaksiran kompetensi secara menyeluruh, skopnya berbeza daripada penilaian prestasi pengajaran simulasi bagi pelatih pengajar TVET-I semasa internship.

Sorotan kajian terdahulu menunjukkan keperluan instrumen penilaian yang lebih khusus bagi mengukur kompetensi pedagogi pelatih pengajar TVET-I berasaskan Metodologi Pengajaran 4-Langkah. Setiap komponen pengajaran perlu diterjemahkan kepada item dan indikator tingkah laku yang jelas, boleh diperhatikan serta dinilai secara konsisten. Keperluan ini selari dengan Ismail et al. (2020) yang mengesahkan model pengukuran kompetensi teknikal pengajar vokasional di institusi latihan kemahiran awam Malaysia, serta Che Amran et al. (2020) yang membangunkan standard kompetensi berasaskan NOSS bagi pensyarah institusi TVET peringkat tinggi. Kedua-dua kajian menunjukkan bahawa kompetensi pengajar TVET boleh dan perlu diukur secara berstruktur, tetapi kedua-duanya tidak menyentuh penilaian pemerhatian pengajaran semasa internship. Justeru, kajian ini memfokuskan analisis keperluan pembangunan instrumen penilaian internship TVET-I dengan mengambil kira kekuatan instrumen sedia ada serta tuntutan kompetensi pedagogi dalam latihan kemahiran semasa.

2.2. Dilema Penilaian Prestasi, Skor Pemotongan Statik & Prinsip Rubrik Analitik

Penilaian Simulasi Mengajar merupakan satu bentuk penilaian prestasi (*performance assessment*) yang membolehkan kompetensi pengajaran diperhatikan melalui pelaksanaan tugas sebenar atau simulasi, sebagaimana amalan pengajaran mikro (*micro-teaching*) dalam pendidikan guru praperkhidmatan (Iliasova et al., 2025). Walau bagaimanapun, penilaian ini berdepan cabaran dari segi tempoh pemerhatian, kos pelaksanaan, keterbatasan bukti prestasi dan ketekalan pemarkahan. Stufflebeam (2001), melalui penelitian terhadap 22 pendekatan penilaian program, membincangkan kelemahan sistem penilaian yang bergantung pada bukti terhad dan prosedur pengukuran yang kurang mantap, termasuk pengalaman pelaksanaan *Kentucky Instructional Results Information System* (KIRIS) di Amerika Syarikat. Dalam perbincangan yang sama, kerja Sanders dan Horn (1994, seperti dipetik dalam Stufflebeam, 2001) menimbulkan persoalan tentang penggunaan skor pemotongan tetap yang boleh mendorong tumpuan pengajaran dan penilaian kepada pencapaian markah lulus minimum. Dalam konteks internship TVET-I, perkara ini menunjukkan kepentingan rubrik analitik dengan deskriptor tingkah laku yang jelas, panduan pemarkahan yang sistematik serta bukti empirikal bagi menyokong kesahan dan kebolehpercayaan keputusan penilaian.

Kelemahan bukti psikometrik dalam instrumen vokasional praktikal telah lama direkodkan. McCray (1982), dengan merujuk Sax (1973), mencatatkan jarangnya penggunaan kaedah statistik bagi membuktikan kebolehpercayaan dan kesahan sampel kerja vokasional; jadual perbandingan sistem sampel kerja komersial dalam manual tersebut mencatatkan nota berulang "*no data available*" bagi bukti kebolehpercayaan atau kesahan sistem seperti Micro-TOWER, JEVS, Valpar #17 dan TOWER. Bagi mengatasi ketidaktepatan skala kadar umum atau senarai semak binari (Terima/Tidak Terima), tinjauan sistematik terhadap 75 kajian oleh Jonsson dan Svingby (2007) menegaskan bahawa rubrik analitik perlu memecahkan prestasi kompleks kepada dimensi tertentu dan menghuraikan perubahan tingkah laku yang boleh diperhatikan merentas aras secara eksplisit, disokong oleh contoh prestasi (*benchmarks*) dan latihan penilai. Dawson (2017) memperincikan elemen reka bentuk rubrik, termasuk kriteria, aras kualiti, definisi kualiti, strategi pemarkahan dan contoh prestasi, manakala meta-analisis Panadero et al. (2023) menunjukkan kesan positif rubrik terhadap prestasi, regulasi kendiri dan efikasi kendiri pelajar. Walau bagaimanapun, kewujudan rubrik tidak menjamin kesahan secara automatik tanpa pengujian empirikal di lapangan (Brookhart, 2018; White & Klette, 2024).

2.3. Paradoks Literasi Penilai & Keperluan Penentukuran Penilaian Pemerhatian

Borang penilaian atau rubrik pemerhatian yang dibangunkan secara sistematik belum tentu menghasilkan keputusan yang konsisten sekiranya penilai mempunyai tahap pengetahuan penilaian yang berbeza atau berdepan beban tugas yang tinggi. Tinjauan Noll (1977, seperti dipetik dalam McCray, 1982) melaporkan bahawa 21% penilai vokasional tidak pernah menerima latihan khusus dalam bidang penilaian, manakala 21% lagi hanya mengikuti latihan berbentuk bengkel. Walaupun data ini sudah lama, ia menunjukkan bahawa isu kesediaan penilai dalam konteks vokasional bukanlah perkara baharu.

Kepentingan kompetensi penilaian turut dibincangkan oleh Xu dan Brown (2016) yang mensintesis 100 kajian bagi membina kerangka *Teacher Assessment Literacy in Practice* (TALiP). Sebelum itu, Alkharusi et al. (2012) yang mengkaji 165 orang guru dalam perkhidmatan di Muscat, Oman, mendapati wujud jurang antara persepsi kecekapan penilaian

dengan pengetahuan sebenar guru. Walaupun 73.3% responden menganggap diri mereka cekap atau sangat cekap dalam penilaian (M = 3.68, SP = 0.44, α = .93), termasuk penilaian prestasi (M = 3.70, SP = 0.50, α = .72), purata pencapaian mereka dalam *Teacher Assessment Literacy Questionnaire* (TALQ) hanyalah 12.42 daripada 32 markah (SP = 3.34, KR-20 = .62), dan 82 orang guru memperoleh 12 markah atau kurang. Dapatan ini menunjukkan bahawa keyakinan penilai tidak semestinya mencerminkan penguasaan pengetahuan penilaian sebenar. Dalam konteks internship TVET-I, keadaan ini mengukuhkan keperluan instrumen yang mempunyai indikator tingkah laku dan panduan penskoran yang jelas bagi membantu penilai membuat pertimbangan secara konsisten.

Kajian yang sama turut mendapati bahawa beban mengajar mingguan mempunyai hubungan negatif yang signifikan dengan persepsi kecekapan guru dalam membangunkan dan memberikan skor penilaian prestasi ($r = -.253$, $p < .01$) serta kecekapan penilaian keseluruhan ($r = -.210$, $p < .01$). Sebaliknya, guru yang mengikuti latihan dalam perkhidmatan menunjukkan tahap pengetahuan penilaian yang lebih tinggi secara signifikan ($F = 7.102$, $p = .009$, separa $\eta^2 = .045$). Dapatan ini menunjukkan bahawa latihan penilaian dan pengurusan beban tugas wajar diberi perhatian dalam usaha meningkatkan kualiti pelaksanaan penilaian.

Walaupun kajian Alkharusi et al. (2012) melibatkan guru sekolah di Oman dan dapatannya tidak boleh digeneralisasikan secara langsung kepada penilai internship TVET-I di Malaysia, kajian tersebut memberikan asas empirikal untuk meneliti faktor penilai dalam pengukuran kompetensi pengajaran. Literatur pemerhatian pengajaran yang lebih terkini mengukuhkan hujah ini. Johnson et al. (2020) mendapati penilai sukar memahami dan mengaplikasikan item rubrik berinferens tinggi secara tepat dan konsisten, manakala White (2018) berhujah bahawa skor pemerhatian hanya sah jika ketepatan penilai diukur berbanding piawaian prestasi penilai yang eksplisit. Sehubungan itu, pembangunan instrumen penilaian internship TVET-I perlu mempertimbangkan deskriptor pemarkahan yang jelas, manual penskoran yang terperinci serta latihan penyelarasan pemarkahan antara penilai (*scorer calibration*). Keperluan ini penting bagi menyokong penggunaan instrumen secara konsisten oleh Penyelia dan PPL, di samping menyediakan asas untuk menilai kebolehpercayaan antara penilai dalam fasa kajian empirikal seterusnya.

2.4. Kerangka DDR Jenis 1 & Kewajipan Pengujian Semula Kesahan Skor

Richey dan Klein (2007) membahagikan DDR kepada dua kluster utama: penyelidikan produk dan alat (Jenis 1) yang meneliti kitaran analisis, reka bentuk, pembangunan dan penilaian sesuatu alat dalam persekitaran kerja sebenar, dan penyelidikan model (Jenis 2) yang memfokuskan pembinaan serta pengesahan model proses reka bentuk yang lebih umum. Kajian utama ditempatkan dalam DDR Jenis 1 kerana matlamatnya ialah membina serta mengesahkan alat pengukuran prestasi pengajaran, dan bukannya model makro sektor TVET. Sebagai contoh gandingan semakan dokumen pengajaran dengan pemerhatian masa terhad dalam tradisi DDR, Richey dan Klein (2007) mengetengahkan kajian Taylor dan Ellis (1991) terhadap 100 kursus latihan Tentera Laut Amerika Syarikat yang menggabungkan semakan lima langkah ke atas dokumen objektif dan item ujian dengan pemerhatian bilik darjah selama 30 minit untuk menilai keselarasan penyampaian.

Kajian ini juga berpandukan *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA] et al., 2014), yang menegaskan bahawa kesahan merujuk sokongan bukti terhadap tafsiran dan tujuan penggunaan skor, bukan sifat borang secara terasing. Kane (2013) menghuraikan prinsip ini sebagai hujah kesahan (*validity argument*) yang merangkumi inferens pemarkahan, generalisasi, ekstrapolasi dan keputusan; Bell et al. (2012) pula mengaplikasikannya kepada protokol pemerhatian bilik darjah dan menunjukkan bahawa latihan dan penentukuran penilai merupakan pautan kritikal dalam hujah tersebut. Creswell dan Creswell (2018) turut mengingatkan bahawa instrumen baharu atau yang diubah suai perlu diuji semula kesahan dan kebolehpercayaannya, dengan nilai ketekalan dalaman Cronbach's alpha antara 0.70 hingga 0.90 lazimnya disyorkan (Tavakol & Dennick, 2011).

3. Metodologi Analisis Dokumentasi

Kajian ini menggunakan reka bentuk analisis dokumentasi kualitatif dan perbandingan dokumen sebagai komponen teras analisis keperluan dalam DDR Jenis 1 (Richey & Klein, 2007). Analisis dokumen ialah prosedur sistematik untuk mencari, memilih, menilai dan mensintesis data daripada dokumen, dan amat sesuai bagi kajian intensif terhadap dasar atau program serta sebagai sumber triangulasi (Bowen, 2009). Hujah jurang pula disusun mengikut model kekurangan (*deficiencies model*) yang dihuraikan oleh Creswell dan Creswell (2018). Pendekatan ini dipilih kerana sistem latihan kemahiran bertauliah dikawal secara langsung oleh standard pekerjaan kebangsaan, garis panduan agensi dan borang pemarkahan rasmi. Oleh itu, bukti keperluan pembangunan instrumen baharu wajar disusun terlebih dahulu melalui semakan silang dokumentasi rasmi sebelum temu bual lapangan atau reka bentuk prototaip dijalankan.

3.1. Himpunan Dokumen yang Dianalisis

Bahan analisis merangkumi dokumen dasar peringkat kebangsaan dan panduan operasi rasmi, borang penilaian yang sedang digunakan, rekod penempatan latihan, serta literatur ilmiah dan teks metodologi yang relevan (JPK, 2021, 2023, 2024; Pusat Latihan Pengajar dan Kemahiran Lanjutan [CIAST], 2025, 2026). Bagi memelihara ketelusan jejak kajian dan memudahkan semakan semula, Jadual 1 menyenaraikan kategori dokumen primer dan sekunder yang dianalisis mengikut kod atau rujukan, edisi atau tarikh, serta peranan analitiknya dalam membina justifikasi kajian utama.

Jadual 1. Bahan Dokumentasi Berkaitan

Kategori Dokumen	Nama Dokumen & Kod / Rujukan	Tarikh / Edisi	Peranan Analitik
Standard Pekerjaan Kebangsaan (NOSS)	NOSS TVET Instruction (TVET-I) Level 4 (P853-002-4:2021)	2021	Menentukan rasional semakan VTO, struktur enam CU (C01–C06), aktiviti kerja (WA), *4-steps pattern*, *Law of Learning*, serta pembezaan teori dan amali
Panduan Kurikulum & Bahan	Panduan Pembangunan Bahan Pengajaran dan Pembelajaran (BPdP) Edisi 2024 (JPK)	Kemas kini 11 Mac 2025 (kuat kuasa 1 Jan 2025)	Menentukan perubahan parameter SKM Tahap 3 berbanding DKM Tahap 4 (Jadual 2, 3, 15, 30), nisbah masa 3:2:1(PMT) vs. 1:3:2(PMA), serta format senarai semak kerja
Panduan Operasi Lawatan	Panduan Tugas dan Peranan Pegawai Penilai Semasa Lawatan Internship Program Pengajaran TVET (TVET-I) (P853-002-4:2021)(CIAST)	Edisi 1, 15 April 2025	Menentukan aliran kerja PP CIAST/CSC, Simulasi Mengajar 30 Minit, markah lulus 60%, ulang nilai, bimbingan maklum balas, dan rujukan Seksyen 5.0
Borang Penilaian Rasmi	JPK/P853/B_PP, B_PPL, C, D (PENYELIA), D_PP, JPK/P853/D_PPL (Latihan Internship TVET I)	Mac 2023 (hlm. 19–33)	Semakan fizikal item demi item (43 item Borang D; 20 item B dan C), skala 0-3, wajaran dimensi, pertindihan WIM dan pengiraan markah akhir
Dokumen Pentadbiran & Rekod Penempatan	Senarai Semak Intern, Kriteria PB Internship dan Penyelia, JPK/P853/G (Laporan Kehadiran Bulanan), Surat Penempatan PB CIAST.PPK/UMV600-3/1/6(16) & Lampiran 1 Jadual Internship TVET-I Sesi Januari 2026	2026 (Surat: 29/7/2026)	Menentukan syarat 720 jam internship, 300 jam mengajar, kelayakan penyelia PB, skala operasi 398 rekod penempatan, anomali rekod, dan rujukan kod P853-002-4:2024
Himpunan Literatur Empirik & Metodologi Terarah	Bidin et al. (2020); Zulkifli (2019); Alkharusi et al. (2012); Stufflebeam (2001); McCray (1982); Richey & Klein (2007); Creswell & Creswell (2018); AERA et al. (2014); Jonsson & Svingby (2007); Shulman (1986)	1982–2020 (himpunan rujukan awal)	Membekalkan kerangka teori 4-Langkah & PCK, prinsip rubrik analitik, piawaian kesahan dan kebolehpercayaan, bukti prinsip literasi penilai, serta metodologi DDR (Sarkar et al., 2024)

3.2. Prosedur Analisis & Triangulasi Silang Dokumen

Analisis dokumentasi dilaksanakan melalui empat peringkat berstruktur yang menggabungkan pengecaman pola kurikulum dengan semakan fizikal instrumen:

1. **Pengekstrakan Struktur dan Parameter Dasar**: Mengekstrak klausa rasmi daripada P853-002-4:2021, Panduan Pembangunan Bahan Pengajaran dan Pembelajaran (BPdP) Edisi 2024 dan Panduan Tugas dan Peranan Pegawai Penilai Semasa Lawatan Internship (Panduan PPI) 2025 berhubung peralihan SKM Tahap 3 kepada DKM Tahap 4, pecahan penilaian COPTPA (*Code of Practice for TVET Programme Accreditation*), julat masa bersemuka, nisbah 3:2:1 dan 1:3:2, serta prosedur Simulasi Mengajar 30 Minit dan penilaian semula (ulang nilai) menggunakan rakaman digital;
2. **Pemeriksaan Fizikal dan Perbandingan Instrumen Rasmi:** Menyemak secara langsung kesemua muka surat enam borang rasmi (B_PP, B_PPL, C, D (PENYELIA), D_PP, D_PPL) untuk mengenal pasti bilangan item, skor mentah maksimum setiap bahagian, skala pemarkahan, ketiadaan atau kewujudan deskriptor tingkah laku, kod pengecualian, formula pengiraan markah, serta pertindihan kandungan antara Borang B dengan Borang D;
3. **Triangulasi Silang dan Pengesanan Percanggahan:** Membandingkan ketetapan kurikulum (NOSS dan BPdP) terhadap operasi lapangan (PPI, Senarai Semak, surat penempatan 2026, Jadual Jan 2026) serta borang penilaian (Borang D) bagi mengesan perubahan versi kod program, ketakselarasan tugas, konstruk dan rekod anomali penempatan; dan
4. **Pemetaan Konstruk dan Penetapan Sempadan Kajian Utama:** Memetakan kesemua 43 item pada JPK/P853/D satu per satu kepada struktur 4-Langkah (Persediaan, Penyampaian, Penggunaan dan Pengesahan), mengasingkan item bukan penyampaian (Item E7 – *WIM*), membezakan indikator khusus teori berbanding amali, serta memisahkan pematuhan pentadbiran daripada skor kompetensi mengajar.

Bagi memastikan kebolehpercayaan tafsiran, setiap dapatan disemak secara silang merentas sekurang-kurangnya dua dokumen primer (triangulasi dokumen), seperti yang disarankan oleh Bowen (2009). Sempadan bukti juga dinyatakan secara telus, termasuk perbezaan antara pengakuan operasi agensi dengan bukti statistik, serta kedudukan kajian luar sebagai bukti prinsip pengukuran dan bukan bukti tentang borang TVET-I. Tafsiran tidak diekstrapolasi melampaui kandungan dokumen sumber.

4. Dapatan Analisis Dokumentasi: Diagnosis Jurang & Bukti Keperluan

Analisis silang dokumen mengenal pasti lima dapatan utama yang menunjukkan jurang pengukuran dan keperluan pembangunan instrumen penilaian internship TVET-I. Keperluan ini selari dengan Mansor et al. (2024) yang melaporkan ketiadaan kerangka penilaian prestasi kerja yang komprehensif bagi graduan pelatih pengajar TVET. Walaupun skopnya berbeza, kedua-dua kajian menunjukkan kepentingan instrumen penilaian yang sistematik, khususnya bagi menilai kompetensi pedagogi pelatih internship TVET-I berasaskan Metodologi Pengajaran 4-Langkah.

4.1. Transformasi Parameter Latihan & Penilaian, VTO Tahap 3 kepada TVET-I Tahap 4

Analisis terhadap NOSS P853-002-4:2021 dan Panduan Pembangunan BPdP Edisi 2024 (berkuat kuasa mulai 1 Januari 2025 dan membatalkan Panduan Pembangunan WIM Edisi 2020) menunjukkan bahawa peralihan daripada VTO Tahap 3 kepada TVET-I Tahap 4 melibatkan perubahan struktur kelayakan dan reka bentuk penilaian yang substantif (JPK, 2021, 2024). Seksyen 1.2 NOSS tersebut menyatakan bahawa standard Tahap 4 dibangunkan sebagai semakan terhadap I-031-3:2014, sekali gus mengemukakan cadangan pelupusan (*proposed obsolescence*) VTO Tahap 3, menaik taraf kelayakan pengajar TVET kepada peringkat Diploma sebagai lesen mengajar TVET serta mewajibkan penempatan internship di PB. Dari segi pemberat latihan teras, dua unit kompetensi yang berkait secara langsung dengan pelaksanaan pengajaran, iaitu C01 (*Instructional Delivery Implementation*, 20%: WA1 20%, WA2 30%, WA3 20%, WA4 20%, WA5 10%) dan C03 (*Competency Based Training Instruction*, 20%: WA1 30%, WA2 30%, WA3–WA5 40%), menyumbang 40% daripada keseluruhan pemberat teras NOSS, disokong oleh pembangunan WIM dalam C02 (15%). Jadual 2 membandingkan parameter antara kerangka terdahulu dengan kerangka semasa berdasarkan dokumen rasmi yang disemak.

Jadual 2. Perubahan Parameter Penilaian VTO & TVET-I

Dimensi Perbandingan	Kerangka Terdahulu (SKM Tahap 3 / VTO I-031-3:2014)	Kerangka Semasa (DKM Tahap 4 / TVET-I P853-002-4:2021)	Implikasi Justifikasi kepada Kajian Lanjutan
Taraf Kelayakan & Prasyarat	Sijil Kemahiran Malaysia (SKM) Tahap 3; cadangan pelupusan *(proposed obsolescence)* dinyatakan dalam NOSS 2021	Diploma Kemahiran Malaysia (DKM) Tahap 4 sebagai lesen mengajar TVET; prasyarat kemasukan sekurang-kurangnya SKM 3/sijil Tahap 3 (Seksyen 1.6)	Instrumen mesti menaksir kompetensi pengajaran kompleks peringkat Diploma dan bukan sekadar rutin penyampaian asas
Pecahan Penilaian COPTPA (Jadual 15 BPdP)	30% Pengetahuan: 70% Prestasi	40% Pengetahuan: 60% Prestasi	Menuntut keseimbangan penilaian & pentaksiran antara penguasaan konsep teori dengan penyeliaan kemahiran amali
Julat Masa Bersemuka Modul Teras (Jadual 30 BPdP)	70–90% bersemuka (5–20% kendiri; 3–20% penilaian)	60–70% bersemuka (10–30% kendiri; 10–30% penilaian)	Penurunan julat masa bersemuka meningkatkan tuntutan kecekapan penyampaian dan integrasi digital (C04)
Latihan Industri / Internship Wajib	Latihan Industri standard Tahap 3 ialah 240 jam (Jadual 2 BPdP); borang pengajaran terdahulu JPK/I031/3D-1	Latihan Industri Tahap 4 ialah 320 jam (Jadual 2 BPdP); khusus TVET-I mewajibkan 720 jam internship dengan minimum 300 jam mengajar	Memerlukan instrumen penilaian & pentaksiran yang dibangunkan dan diuji khusus dalam konteks penempatan internship di PB
Struktur Kompetensi & Bukti Psikometrik	5 kriteria dalam JPK/I031/3D-1 (Bidin et al., 2020); tiada skor empirikal atau data psikometrik dilaporkan	6 CU teras (C01 20%, C02 15%, C03 20%, C04 15%, C05 20%, C06 10%); borang JPK/P853/D menggantikan VTO tanpa data psikometrik	Kesahan instrumen lama tidak boleh diandaikan berterusan; pengujian semula kesahan dan kebolehpercayaan wajib dijalankan (Creswell & Creswell, 2018)

Peralihan kepada DKM Tahap 4 turut mengekalkan *4-steps pattern* dan falsafah *Laws of Learning* sebagai teras C01 WA4 serta membezakan *Bloom's Taxonomy* (teori) daripada *Simpson's Taxonomy*, *shop talk* dan penyeliaan bengkel (amali) dalam C03 WA1–WA2, di samping memasukkan pembangunan PMT/PMA dalam C02. Istilah *Conformation* dalam jadual asal NOSS difahami sebagai merujuk langkah Pengesahan (*Confirmation*).

4.2. Pemeriksaan Struktur Borang Penilaian Rasmi JPK/P853

Pemeriksaan fizikal ke atas enam borang rasmi dalam Latihan Internship TVET-I (P853-002-4:2021 – CIAST – Mac 2023) menunjukkan bahawa instrumen TVET-I sebenarnya telah dibina, tetapi pada masa yang sama mendedahkan bukti

objektif mengapa instrumen sedia ada belum memenuhi ciri rubrik analitik dan keperluan kebolehpercayaan pengukuran. Ketiga-tiga borang Simulasi Mengajar, iaitu JPK/P853/D (PENYELIA), JPK/P853/D_PP dan JPK/P853/D_PPL, memuatkan 43 item yang sama merentas lima bahagian dengan skor mentah maksimum masing-masing 30 (Bahagian A: 10 item), 39 (Bahagian B: 13 item), 15 (Bahagian C: 5 item), 15 (Bahagian D: 5 item) dan 30 (Bahagian E: 10 item), menghasilkan jumlah skor mentah maksimum 129. Borang B_PP, B_PPL dan C pula masing-masing mengandungi 20 item dengan skor mentah maksimum 60. Jadual 3 merumuskan struktur serta kelemahan pengukuran yang ditemui pada keenam-enam borang tersebut.

Dapatan dalam Jadual 3 mengukuhkan tiga hujah utama bagi pembangunan instrumen baharu. Pertama, kesemua 43 item Simulasi Mengajar hanya menggunakan label skala umum (0 – Sangat Kurang Memuaskan, 1 – Kurang Memuaskan, 2 – Memuaskan, 3 – Baik) tanpa huraian perlakuan yang membezakan satu aras daripada aras lain; menurut Jonsson dan Svingby (2007), keadaan ini membuka ruang kepada tafsiran subjektif penilai serta melemahkan maklum balas diagnostik kepada pelatih. Kedua, kemasukan Item 7 dalam Bahagian E ("Menyediakan kandungan dan format WIM yang berkualiti") mengekalkan struktur warisan VTO (JPK/I031/3D-1) dan mencampuradukkan kualiti dokumentasi WIM ke dalam skor prestasi pengajaran, sedangkan dokumen WIM telah dinilai dalam borang B_PP (komponen 15%) dan B_PPL (komponen 30%). Ketiga, ketiadaan pilihan 'Tidak Berkenaan' dan 'Bukti Tidak Mencukupi' menjejaskan ketepatan tafsiran skor apabila sesuatu perlakuan tidak relevan atau tidak berpeluang diperhatikan semasa simulasi. Selain itu, Jadual 3 menunjukkan bahawa item dan wajaran yang sama digunakan bagi sesi teori dan amali (dibincangkan dalam Bahagian 4.3), dan bahawa pemberat akhir bagi borang D_PP (10%) dan D_PPL (30%) berbeza walaupun instrumennya sama; justifikasi perbezaan ini perlu dijelaskan dalam kajian lanjutan.

Jadual 3. Kelemahan Pengukuran Hasil Triangulasi Silang Dokumen

Kod Borang Rasmi	Penilai & Masa Pelaksanaan	Struktur Item, Skala & Rumus Pemarkahan	Kelemahan Struktur & Risiko Pengukuran Dikesan
JPK/P853/B_PP (hlm. 19–20)	Pegawai Penilai (PP CIAST/CSC) semasa Lawatan Internship	20 item (Set WIM Teori 5, WIM Amali 4, PKK (Penilaian Kerja Kursus) Pengetahuan 3, PKK Prestasi 3, ABM 3, Keseluruhan 2); skala 0–3; markah (/ 60) × 15%; lulus 60%	Label 0–3 tanpa deskriptor tingkah laku; menilai format dan kualiti 3 set WIM, PKK dan ABM secara berasingan
JPK/P853/B_PPL (hlm. 21–23)	Pegawai Pengesah Luaran (PPL) pada Verifikasi 720 jam	20 item identikal dengan B_PP; skala 0–3; markah (/ 60) × 30%; jumlah akhir gabungan PP (15%) + PPL (30%)	Label 0–3 tanpa deskriptor tingkah laku; tiada panduan pembezaan aras kualiti antara skor 1, 2, dan 3
JPK/P853/C (hlm. 24)	Penyelia PB (Penilaian 1 sebelum lawatan; Penilaian 2 sebelum PPL); disahkan Lulus/ Gagal oleh PP	20 item sikap, disiplin, keselamatan & pengurusan kelas/bengkel; skala 0–3; Penilaian 1 (/ 60) × 5%, Penilaian 2 (/ 60) × 10%; lulus 60%	Label 0–3 tanpa deskriptor; menggabungkan kelakuan am tempat kerja dengan item pengajaran umum tanpa definisi tingkah laku
JPK/P853/D (PENYELIA) (hlm. 25–27)	Penyelia PB sebelum Lawatan Internship	43 item (skor maksimum 129) dalam 5 dimensi: Persediaan (10, 15%), Penyampaian (13, 35%), Penggunaan (5, 20%), Pengesahan (5, 10%), Teknik Umum Mengajar (10, 20%); skala 0–3; rumus (f / x) × w; berhenti pada MARKAH %	Tiada deskriptor aras 0–3; tiada kod Tidak Berkenaan / Bukti Tidak Mencukupi; Item 7 Bahagian E bertindih dengan WIM (B_PP); tiada pemberat akhir program
JPK/P853/D_PP (hlm. 28–30)	PP CIAST/CSC semasa Lawatan Internship (Simulasi 30 minit)	43 item dan rumus (f / x) × w yang identikal dengan D (PENYELIA); pengiraan akhir JUMLAH MARKAH (y × 10%)	Kesemua kelemahan sama seperti di atas (tiada deskriptor, tiada kod pengecualian, Item 7 WIM, ralat cetak BAHAGIAN D, item & wajaran sama untuk Teori T dan Amali A)
JPK/P853/D_PPL (hlm. 31–33)	PPL pada Verifikasi Akhir 720 jam (Simulasi 30 minit)	43 item dan rumus (f / x) × w yang identikal; pengiraan akhir JUMLAH MARKAH (y × 30%)	Kesemua kelemahan sama seperti D_PP

4.3. Paradoks Pedagogi Nisbah PMT/PMA Simulasi Mengajar & Ulang Nilai

Semakan silang antara NOSS P853-002-4:2021, Panduan Pembangunan BPdP Edisi 2024 dan Panduan PPI (2025) menunjukkan perbezaan keperluan pedagogi antara sesi teori dan amali yang perlu diberi perhatian dalam pelaksanaan

Simulasi Mengajar 30 minit dan penilaian semula (ulang nilai). Berdasarkan Seksyen 7.1 dan 7.4 Panduan Pembangunan BPdP, pembahagian masa bagi langkah Penyampaian, Penggunaan dan Pengesahan selepas langkah Persediaan adalah berbeza antara Pelan Mengajar Teori (PMT) dengan Pelan Mengajar Amali (PMA), sebagaimana ditunjukkan dalam Jadual 4.

Perbezaan ini memberikan implikasi terhadap peluang pemerhatian dan penilaian kompetensi pengajaran dalam tempoh simulasi yang terhad. Sesi teori lebih menekankan penerangan konsep, penyoalan dan pengukuhan kefahaman, manakala sesi amali memerlukan demonstrasi, latihan kendiri, pemantauan keselamatan serta bimbingan dan pembetulan kesilapan pelatih. Dalam tempoh 30 minit, tidak semua tingkah laku pengajaran berpeluang diperhatikan, khususnya apabila sesuatu kesilapan pelatih tidak berlaku atau tugasan amali memerlukan masa yang lebih panjang. Oleh itu, pemilihan kandungan dan skop tugasan simulasi perlu membolehkan keempat-empat langkah pengajaran dilaksanakan, di samping mempertimbangkan indikator yang bersesuaian serta kaedah merekod bukti pemerhatian yang mencukupi. Pertimbangan yang sama relevan bagi penilaian semula menggunakan rakaman digital, kerana sudut rakaman mungkin mengehadkan pemerhatian terhadap interaksi atau tindakan tertentu.

Walau bagaimanapun, peruntukan masa dalam pelan sesi, iaitu nisbah 3:2:1 bagi teori dan 1:3:2 bagi amali, tidak semestinya menentukan wajaran markah penilaian mengikut nisbah yang sama. Sebaliknya, perbezaan sifat dan tuntutan pengajaran antara kedua-dua sesi menunjukkan bahawa kesesuaian 43 item dan wajaran penilaian yang seragam perlu diteliti secara empirikal. Justeru, instrumen penilaian internship TVET-I dalam kajian lanjutan perlu mempertimbangkan indikator bersama dan indikator khusus bagi sesi teori serta amali, dengan mengekalkan keselarasan terhadap konstruk Metodologi Pengajaran 4-Langkah.

Jadual 4. Perbandingan Penilaian PMT dan PMA

Aspek Perbandingan	Pelan Mengajar Teori (PMT) & Sesi Teori	Pelan Mengajar Amali (PMA) & Sesi Amali	Ketidaksejajaran pada Borang D & Risiko Simulasi 30 Minit
Fokus Taksonomi & Kandungan (NOSS C03)	*Related Knowledge*; berteraskan *Bloom's Taxonomy*; merujuk Kertas Penerangan (KP) dan Kertas Tugasan (KT) sebagai PKK Pengetahuan	*Related Skills*; berteraskan *Simpson's Taxonomy*, teknik shop talk, demonstrasi langkah kerja, penyeliaan bengkel, sikap, keselamatan & persekitaran; merujuk Kertas Kerja (KK)	Walaupun Borang D menyediakan pilihan T: Teori dan A: Amali, kesemua 43 item dan wajaran (15% / 35% / 20% / 10% / 20%) adalah identikal
Nisbah Masa Rasmi BPdP (Selepas Persediaan)	**3: 2: 1** (Contoh rasmi 2 jam / 120 minit: Persediaan 10 minit, Penyampaian 55 minit, Penggunaan 35 minit, dan Pengesahan 20 minit)	**1: 3: 2** (Contoh rasmi 6 jam / 360 minit: Persediaan 15 minit, Penyampaian 1 jam, Penggunaan 2 jam 50 minit, dan Pengesahan 1 jam 55 minit)	Wajaran Borang D meletakkan Penyampaian tertinggi (35%) berbanding Penggunaan (20%) dan Pengesahan (10%) bagi kedua-dua mod tanpa pengujian kesetaraan item
Pelaksanaan Kitaran 4-Langkah dalam 30 Minit	Memerlukan penstrukturan sesi mikro teori secara berkadaran mengikut nisbah 3:2:1 selepas Persediaan supaya keempat-empat langkah sempat berlaku	Memerlukan pemilihan sub-tugasan kerja yang munasabah serta penstrukturan mikro amali secara berkadaran mengikut nisbah 1:3:2 hingga fasa Pengesahan	Ketiadaan protokol pemilihan segmen sub-tugasan amali menyebabkan sesi bengkel 30 minit berisiko terhenti sebelum Pengesahan atau gagal menilai interaksi
Kerentanan Item Bergantung Interaksi (C2, Ulang Nilai)	Item seperti Kesilapan pelatih diperbetulkan segera (C2) bergantung pada wujudnya kekeliruan atau jawapan salah pelatih	Item seperti Kesilapan pelatih diperbetulkan segera (C2) dan penyeliaan keselamatan bergantung pada kesilapan kerja *hands-on* dan pemerhatian penilai.	Tanpa kod Bukti Tidak Mencukupi, ketiadaan peluang kesilapan dalam 30 minit atau had rakaman ulang nilai berisiko dikodkan sebagai skor 0 atau 1

4.4. Pemetaan Lengkap 43 Item Borang D ke Struktur Konstruk 4-Langkah

Untuk menunjukkan kebolehlaksanaan penstrukturan semula instrumen, kesemua 43 item daripada JPK/P853/D_PP, D (PENYELIA) dan D_PPL dipetakan satu demi satu kepada empat konstruk teras Metodologi Pengajaran 4-Langkah: (1) Persediaan Pembelajaran; (2) Penyampaian Pengajaran; (3) Penggunaan dan Bimbingan; dan (4) Pengesahan dan Penutupan. Dalam penstrukturan semula ini, Item E7 (Menyediakan kandungan dan format WIM yang berkualiti) dikeluarkan daripada skor pemerhatian mengajar dan dialihkan kepada semakan konteks keselarasan dokumen sesi (PMT/PMA), manakala sembilan item pedagogi lain daripada Bahagian E diserap ke dalam empat langkah pengajaran

tersebut. Jadual 5 memaparkan matriks pemetaan lengkap berserta sandaran standard, kebolehgunaan sesi dan keperluan kod pengecualian. Pemetaan ini ialah cadangan awal pengkaji dan akan disahkan oleh panel pakar dalam Fasa II.

Nota: Dalam Jadual 5, 'Langkah 1–4' merujuk empat langkah Metodologi Pengajaran 4-Langkah (Persediaan, Penyampaian, Penggunaan dan Pengesahan), manakala 'Langkah' bernombor dalam rujukan BPdP (Rajah 1C–1M) merujuk langkah dalam rajah alir BPdP.

Jadual 5. Matriks Pemetaan Struktur Metodologi Pengajaran 4-Langkah

Item asal JPK/P853/D (Borang D, Mac 2023)	**Konstruk baharu** (4-Langkah)	**Sandaran standard** (NOSS / BPdP 2024)	**Kebolehgunaan sesi dan keperluan kod pengecualian**
A1 (Sebelum): Penampilan diri yang kemas	Langkah 1: Persediaan Pembelajaran	C01 WA4 *(Personal appearance)*; BPdP etika pengajar	Bersama (Teori & Amali); amali merangkumi pematuhan pakaian keselamatan/PPE
A2 (Sebelum): Mewujudkan persekitaran yang selesa	Langkah 1: Persediaan Pembelajaran	C01 WA2 & WA3; BPdP Rajah 1C/1I	Bersama (Teori: susun atur kelas; Amali: keselesaan & keselamatan ruang kerja)
A3 (Sebelum): Bersedia dengan bahan pembelajaran	Langkah 1: Persediaan Pembelajaran	C01 WA2 *(Training material)*; C03 WA1–WA2 (IS/WS)	Bersama (KP bagi Teori; KK bagi Amali); semak keselarasan dengan PMT/PMA
A4 (Sebelum): Memastikan peralatan yang cukup dan berfungsi dengan baik	Langkah 1: Persediaan Pembelajaran	C01 WA2 *(Training facilities)*; C03 WA2 (TEM)	Bersama (Teori: AVA/projektor/komputer; Amali: TEM bengkel & keselamatan)
A5 (Semasa 1): Memberi salam dan Semak kehadiran	Langkah 1: Persediaan Pembelajaran	BPdP Rajah 1C & 1I (Langkah 1 & 2 Persediaan)	Bersama (Teori & Amali)
A6 (Semasa 2): Selesa dan timbulkan minat pelatih	Langkah 1: Persediaan Pembelajaran	C01 WA3 & WA4 *(Law of Readiness)*; BPdP Langkah 1, 3–5	Bersama; menggabungkan keselesaan, disiplin, motivasi & rangsangan minat
A7 (Semasa 3): Tajuk diperkenalkan dengan jelas	Langkah 1: Persediaan Pembelajaran	BPdP Rajah 1C & 1I Langkah 6 (nyata & tulis di papan putih)	Bersama; perlu dinyatakan dan dipaparkan mengikut format tajuk PMT/PMA
A8 (Semasa 4): Objektif/ hasil diperjelaskan	Langkah 1: Persediaan Pembelajaran	C01 WA3 (3.4); BPdP Langkah 7 (rujuk Tujuan Pengajaran)	Bersama; merujuk kepada 2–5 Tujuan Pengajaran dalam PMT/PMA
A9 (Semasa 5): Menjelaskan faedah / kepentingan pelajaran	Langkah 1: Persediaan Pembelajaran	BPdP Rajah 1C & 1I Langkah 8	Bersama (Teori & Amali)
A10 (Semasa 6): Siasat pengetahuan / pengalaman pelatih	Langkah 1: Persediaan Pembelajaran	C01 WA3 (3.3); BPdP Langkah 5 & 9 *(soal jawab awal)*	Bersama (Teori & Amali)
B1: Kaitkan pelajaran lepas dengan pelajaran baharu	Langkah 2: Penyampaian Pengajaran *(atau jambatan Langkah 1–2)*	C01 WA3 (3.3); *Law of Readiness/ Primacy*	Bersama; kedudukan transisi disahkan bersama pakar dalam Fasa II
B2: Kandungan disampaikan dalam susunan yang logik	Langkah 2: Penyampaian Pengajaran	C01 WA4 (4.2); BPdP Seksyen 4.1	Bersama (Teori: urutan konsep; Amali: urutan langkah kerja)
B3: Pelajaran disampaikan menggunakan kaedah yang bersesuaian	Langkah 2: Penyampaian Pengajaran	C01 WA4 *(Lecture, demo, discussion)*; C03 WA1–WA2	Dipecah deskriptor: penerangan/soal jawab (Teori) vs. shop talk & demonstrasi langkah kerja (Amali)
B4: Penyampaian yang terang dan jelas	Langkah 2: Penyampaian Pengajaran	C01 WA4 *(Fundamental characteristic of voice)*	Bersama; kejelasan vokal, nada, kelajuan (pace), dan intonasi

Item asal JPK/P853/D (Borang D, Mac 2023)	**Konstruk baharu** (4-Langkah)	**Sandaran standard** (NOSS / BPdP 2024)	**Kebolehgunaan sesi dan keperluan kod pengecualian**
B5: Mengajar ditahap keupayaan pelatih	Langkah 2: Penyampaian Pengajaran	PCK Shulman (1986); C03 WA1 (Bloom) / WA2 (Simpson) (Sarkar et al., 2024)	Bersama; penyesuaian penerangan/bimbingan mengikut tahap kefahaman pelatih
B6: Menekankan perkara yang penting	Langkah 2: Penyampaian Pengajaran	C01 WA4 (4.10 *Importance key point*); *Law of Intensity*	Bersama (Teori: konsep teras; Amali: *key points*, ketepatan & prosedur kritikal)
B7: Menekankan aspek keselamatan	Langkah 2: Penyampaian Pengajaran	C01 WA4 (*Safety 4.1*); BPdP Rajah 1J; C06	Kritikal dalam Amali; dalam Teori boleh wujud sebagai prinsip HSE berkaitan atau ditanda Tidak Berkenaan jika sesi konsep tulen tanpa isu keselamatan
B8: Memberi contoh-contoh yang relevan	Langkah 2: Penyampaian Pengajaran	PCK Shulman (1986); *Law of Intensity* (Sarkar et al., 2024)	Bersama (Teori & Amali)
B9: Melibatkan pelatih dalam aktiviti pembelajaran	Langkah 2: Penyampaian Pengajaran *(jambatan ke Langkah 3)*	C01 WA4 (4.3 *Techniques of training activities*)	Bersama; perincian sempadan dengan Langkah 3 disahkan melalui temu bual kognitif
B10: Mengekalkan *eye contact*	Langkah 2 Penyampaian Pengajaran	C01 WA4 (*Quality of delivery: eye contact*)	Bersama; dalam mod ulang nilai rakaman digital perlukan kod Bukti Tidak Mencukupi jika persembahan tidak jelas
B11: Mempamerkan gesture dan posture yang sesuai	Langkah 2: Penyampaian Pengajaran	C01 WA4 (*Posture, gesture, mannerism*)	Bersama; perlukan kod Bukti Tidak Mencukupi jika persembahan tidak jelas
B12: Menggunakan bahasa standard yang jelas	Langkah 2: Penyampaian Pengajaran	C01 WA4 (*Language, effective communication*)	Bersama (Teori & Amali)
B13: Menghadkan kandungan sewajarnya	Langkah 2: Penyampaian Pengajaran	PR290029 (*curriculum limitation*); BPdP(3:2:1 / 1:3:2)	Bersama; mengelakkan lambakan kandungan dalam had simulasi 30 minit
C1: Pelatih membuat perkara penting	Langkah 3: Penggunaan dan Bimbingan	C01 WA4 (*Law of Exercise*); BPdP Rajah 1E/1L	Dipecah deskriptor: pembacaan/latihan KP(Teori) vs. latihan kerja hands-on merujuk KK & senarai semak (Amali)
C2: Kesilapan pelatih diperbetulkan segera	Langkah 3: Penggunaan dan Bimbingan	PR290029 (*Assistance if mistake happen*); PCK (Sarkar et al., 2024)	Bersama; wajib ada kod Bukti Tidak Mencukupi jika tiada kesilapan pelatih berlaku dalam 30 minit
C3: Tahap kefahaman pelatih disemak	Langkah 3: Penggunaan dan Bimbingan	BPdP Rajah 1E Langkah 4 & Rajah 1L Langkah 7 (soalan lisan)	Bersama (Teori: soalan berkaitan KP; Amali: soalan lisan berkaitan langkah kerja)
C4: Pastikan tujuan pembelajaran sudah dicapai	Langkah 3 *(atau jambatan ke Langkah 4)*	C01 WA4 (4.5 *LO reviewed*) & WA5 (5.1)	Bersama; disemak kedudukannya dalam Fasa II agar tidak bertindih dengan Item D4
C5: Memberi maklumbalas yang membina	Langkah 3: Penggunaan dan Bimbingan	BPdP Rajah 1E/1L (akui jawapan, beri pujian/motivasi)	Bersama (Teori & Amali)
D1: Tugasan dijalankan dalam tempoh yang sesuai	Langkah 4: Pengesahan dan Penutupan	BPdP Rajah 1F/1M (maklum masa mula/tamat; pengurusan masa)	Bersama; merujuk tugasan tanpa bantuan mengikut baki nisbah 3:2:1 / 1:3:2

Item asal JPK/P853/D (Borang D, Mac 2023)	**Konstruk baharu** (4-Langkah)	**Sandaran standard** (NOSS / BPdP 2024)	**Kebolehgunaan sesi dan keperluan kod pengecualian**
D2: Nyatakan dengan terang tugas pelatih	Langkah 4: Pengesahan dan Penutupan	BPdP Rajah 1F (simpan KP) / Rajah 1M (rujuk KK & peluang bertanya)	Dipecah deskriptor: ujian lisan/bertulis/digital (Teori) vs. tugasan amali tanpa bantuan (Amali)
D3: Tugasan berasaskan kandungan pelajaran	Langkah 4: Pengesahan dan Penutupan	BPdP Rajah 1F Langkah 2 & 5; C01 WA5	Bersama; keselarasan soalan/tugasan pengesahan dengan KP/KK sesi
D4: Dapat mengukur tujuan pelajaran	Langkah 4: Pengesahan dan Penutupan	C01 WA5 (5.1 *Learning evidence checked*)	Bersama; menyemak sama ada tugasan tanpa bantuan benar-benar mengukur 2–5 tujuan PMT/PMA
D5: Membuat rumusan pelajaran	Langkah 4: Pengesahan dan Penutupan	C01 WA5 (5.2–5.3); BPdP Rajah 1F/1M (rumusan, faedah, tajuk depan, tutup)	Bersama; termasuk arahan mengemas tempat kerja bagi sesi Amali (Rajah 1M Langkah 5)
E1: Mengikut Pelan Mengajar	**Serap ke Langkah 2** *(dan semakan konteks PMT/PMA)*	C01 WA4 (4.2); C03 WA1–WA2	Bersama; menilai kesetiaan penyampaian sebenar terhadap PMT/PMA yang disediakan
E2: Menunjukkan penguasaan dalam tajuk yang diajar	**Serap ke Langkah 2**	PCK Shulman (1986); C03 WA1–WA2 (Sarkar et al., 2024)	Bersama; ketepatan fakta teknikal dan keyakinan demonstrasi prosedur
E3: Mengawal emosi semasa mengajar	**Serap ke Langkah 2**	C01 WA4 (ketenangan, profesionalisme pengajar)	Bersama; kawalan emosi dan tindak balas profesional sepanjang sesi
E4: Pembahagian masa mengajar yang betul	**Serap ke Langkah 4** *(pengurusan masa keseluruhan)*	C01 WA4 (4.11 *Session plan time management*); BPdP 3:2:1/1:3:2	Bersama; kepatuhan agihan masa berkadaran keempat-empat fasa dalam 30 minit
E5: Mewujudkan komunikasi dua hala	**Serap ke Langkah 3** *(turut relevan di Langkah 1 & 2)*	C01 WA4 (*Verbal & non-verbal*); BPdP Rajah 1C Langkah 4	Bersama; interaksi pengajar–pelatih sepanjang sesi
E6: Menggunakan teknik menyoal yang baik	**Serap ke Langkah 3** *(turut relevan di Langkah 1)*	C01 WA4 (4.7 *Question and answer handling*)	Bersama; kualiti penyoalan menyiasat pengetahuan (Langkah 1) dan menguji kefahaman (Langkah 3)
E7: Menyediakan kandungan dan format WIM yang berkualiti	**DIKELUARKAN** daripada skor pengajaran	Sudah dinilai dalam JPK/P853/B_PP (komponen 15%) dan B_PPL (komponen 30%)	Diganti dengan semakan konteks keselarasan dokumen PMT/PMA bagi sesi simulasi sahaja (elak pertindihan konstruk)
E8: Menggunakan Alat Pandang Dengar / Alat Bantu Mengajar dengan berkesan	**Serap ke Langkah 2**	C01 WA4 (*Handling training aids, AVA*)	Bersama; pengendalian papan putih, slaid, model atau peralatan tunjuk cara
E9: Mengaplikasikan teknik pengajaran dan pembelajaran digital	**Serap ke Langkah 2** *(atau Langkah 4 bagi aplikasi digital)*	C01 WA4 (*Digital training material*); C04; BPdP Rajah 1F Langkah 5	Bersama; perlukan kod Tidak Berkenaan jika konteks bengkel fizikal tidak menuntut aplikasi digital semasa demonstrasi
E10: Pelatih sentiasa mengikuti pelajaran	**Serap ke Langkah 3** *(pemantauan penglibatan)*	C01 WA4 (*Trainees motivation*); *Law of Exercise*	Bersama; pengurusan perhatian dan penglibatan berterusan pelatih dalam aktiviti

4.5. Kepelbagaian Ekosistem Penempatan, Keperluan Pentadbiran & Perubahan Standard

Analisis terhadap dokumen pengurusan dan pelaksanaan internship TVET-I mengenal pasti tiga aspek yang perlu diberi perhatian dalam pembangunan instrumen penilaian pengajaran, iaitu kepelbagaian tempat penempatan, keperluan pentadbiran dan perubahan versi standard. Ketiga-tiga aspek ini penting bagi memastikan pemilihan sampel, penentuan konstruk penilaian dan pelaksanaan kajian utama mengambil kira keadaan sebenar persekitaran latihan kemahiran.

Pertama, kepelbagaian institusi dan penempatan internship. Berdasarkan Jadual Internship TVET-I Sesi Januari 2026, sebanyak 398 rekod penempatan disenaraikan merangkumi Semenanjung Malaysia, Sabah dan Sarawak. Penempatan tersebut diselaraskan oleh CIAST Shah Alam bersama rangkaian Kampus Satelit CIAST (CSC), melibatkan pelbagai institusi latihan awam dan swasta seperti ADTEC, ILP, JMTI, IKBN, IKTBN, IKM/TVET MARA, GIATMARA, Kolej Vokasional (KV), sekolah menengah, pusat kemahiran negeri serta kolej dan akademi swasta. Kepelbagaian ini menggambarkan perbezaan persekitaran latihan dari segi kemudahan, bidang kemahiran, kumpulan pelatih dan amalan pengajaran yang perlu dipertimbangkan dalam pembangunan instrumen penilaian.

Walau bagaimanapun, jadual tersebut turut mengandungi beberapa rekod yang memerlukan pengesahan lanjut, termasuk status ‘BERHENTI’ (bil. 266), ‘TANGGUH INTERNSHIP’ (bil. 79), ‘(*) KIV’ (bil. 32–33), catatan kohort terdahulu bagi Sesi September 2025 (bil. 276, 348 dan 356), serta calon yang sama ditempatkan di dua PB berbeza dengan status ‘TBA’ (bil. 338 dan 357). Terdapat juga kemungkinan kesilapan catatan tahun 2016 pada tarikh lawatan tertentu (bil. 390–391). Oleh itu, jumlah 398 merujuk keseluruhan rekod penempatan yang disenaraikan dan bukan semestinya bilangan pelatih internship aktif. Pengesahan data bersama penyelaras internship perlu dilaksanakan sebelum populasi dan sampel kajian utama ditentukan.

Kedua, keperluan pentadbiran dan penilaian kompetensi pengajaran. Semakan terhadap dokumen Kriteria Pusat Bertauliah Internship dan Penyelia, Senarai Semak Intern serta Borang JPK/P853/G (Laporan Kehadiran Bulanan) menunjukkan bahawa pelaksanaan internship tertakluk kepada beberapa syarat pentadbiran. Antaranya, PB perlu berdaftar dengan JPK, menawarkan program SKM yang bersesuaian dan mempunyai pelatih sepenuh masa. PB juga perlu menyediakan peluang mengajar sekurang-kurangnya 300 jam, termasuk minimum 150 jam (atau 50 jam sebulan) dalam tiga bulan pertama sebelum lawatan penilaian, serta menetapkan jadual latihan lima hari seminggu dengan enam jam sehari. Selain itu, maklumat peribadi pelatih internship tidak boleh digunakan bagi tujuan pengiktirafan dalam MySPIKE tanpa kebenaran. Penyelia Internship yang dilantik pula perlu memiliki SKM dalam bidang berkaitan, pengalaman kerja sekurang-kurangnya dua tahun dan kelayakan sebagai pengajar vokasional.

Bagi pelatih internship, keperluan pentadbiran meliputi penyerahan Borang JPK/P853/I dalam tempoh lima hari bekerja, penghantaran Borang JPK/P853/G selewat-lewatnya pada 4 haribulan berikutnya, kehadiran minimum 80% serta penyediaan Buku Log harian sepanjang tempoh latihan 720 jam. Walaupun semua keperluan tersebut penting bagi memastikan internship dilaksanakan secara teratur, pematuhan pentadbiran perlu dibezakan daripada pencapaian kompetensi pedagogi. Oleh itu, instrumen yang dibangunkan dalam kajian utama perlu memberikan tumpuan kepada prestasi pengajaran yang boleh diperhatikan dan diukur berdasarkan Metodologi Pengajaran 4-Langkah, manakala maklumat pematuhan pentadbiran berfungsi sebagai bukti sokongan dan rujukan konteks.

Ketiga, penyelarasan prosedur, istilah dan versi standard. Dari aspek tadbir urus, Panduan PPI (2025) digunakan sebagai panduan pelaksanaan tugas PP dalam rangkaian CIAST dan CSC. Panduan ini merujuk Perkara 12.2(ii) dalam panduan program JPK bertarikh 1 April 2022 yang menetapkan satu lawatan penilaian internship serta penghantaran Borang JPK/P853/E dalam tempoh satu minggu selepas lawatan, dengan peruntukan dua jam Standard Kualiti Usaha (SKU) bagi setiap pelatih. Seksyen 1.2 panduan tersebut turut mengakui keperluan menyelesaikan isu ketidakseragaman dalam pelaksanaan penilaian internship. Perkara ini menyediakan asas untuk meneliti kejelasan prosedur, keseragaman pemarkahan dan keperluan panduan penilaian yang lebih terperinci.

Selain itu, istilah ‘PP’ digunakan secara berbeza dalam Panduan BPdP (2024), yang merujuk tenaga pengajar atau Pegawai Penilai di PB, berbanding Panduan PPI (2025), yang merujuk Pegawai Penilai CIAST dan CSC. Perbezaan antara PP di PB, PP CIAST dan PP CSC ini perlu diperjelas bagi mengelakkan kekeliruan tentang peranan dan tanggungjawab penilai. Dari segi rujukan standard, beberapa dokumen, termasuk Panduan PPI (April 2025), Borang JPK/P853/G dan borang Latihan Internship Mac 2023, masih merujuk NOSS P853-002-4:2021, sedangkan surat rasmi penempatan CIAST bertarikh 29 Julai 2026 (CIAST.PPK/UMV 600-3/1/6(16)) telah menggunakan kod P853-002-4:2024. Oleh itu, status dan kandungan standard versi 2024 belum jelas dan perlu disahkan sebelum dijadikan rujukan dalam pembangunan instrumen.

Secara keseluruhannya, analisis dokumen ini menunjukkan bahawa pembangunan instrumen penilaian internship TVET-I perlu mengambil kira kepelbagaian konteks penempatan, pemisahan antara keperluan pentadbiran dengan kompetensi pedagogi, serta penyelarasan istilah dan rujukan standard. Penemuan ini menyediakan asas untuk menentukan populasi kajian yang tepat, memperjelas konstruk penilaian dan membangunkan instrumen berasaskan Metodologi Pengajaran 4-Langkah yang sesuai digunakan dalam pelbagai persekitaran latihan kemahiran.

5. Perbincangan: Justifikasi Kajian Lanjutan & Spesifikasi Reka Bentuk Instrumen

Berdasarkan keseluruhan dapatan analisis dokumentasi, kajian lanjutan untuk membangunkan dan mengesahkan instrumen penilaian pengajaran pelatih internship TVET-I berasaskan Metodologi Pengajaran 4-Langkah berdiri di atas tiga teras justifikasi yang saling mengukuhkan, dan bukan sekadar bersandarkan alasan ketiadaan kajian terdahulu.

5.1. Rantaian Justifikasi Kajian Lanjutan Pembangunan Instrumen Penilaian TVET-I

Pertama, dari sudut dasar dan kurikulum, peralihan rasmi daripada VTO Tahap 3 (SKM) kepada TVET-I Tahap 4 (DKM) meningkatkan pecahan penilaian pengetahuan kepada 40% (COPTPA), menurunkan julat masa bersemuka kepada 60–70%, menempatkan 40% daripada pemberat teras NOSS pada unit penyampaian pengajaran (C01 dan C03), memasukkan latihan berdigital (C04) dan mewajibkan latihan internship 720 jam di PB. Perubahan ini menuntut instrumen penilaian pengajaran yang benar-benar mencerminkan kerumitan pedagogi Tahap 4 serta perbezaan mod teori (nisbah masa 3:2:1) dan amali (nisbah masa 1:3:2), bukannya mewarisi struktur lima kriteria era VTO (JPK/I031/3D-1).

Kedua, dari aspek struktur pengukuran, semakan terhadap borang JPK/P853/D_PP, D (PENYELIA) dan D_PPL menunjukkan bahawa walaupun instrumen rasmi telah tersedia, kesemua 43 item menggunakan skala 0–3 tanpa deskriptor tingkah laku. Oleh itu, instrumen tersebut lebih bersifat skala kadar berbanding rubrik penilaian (Brookhart, 2018; Dawson, 2017). Instrumen ini juga belum membezakan keadaan 'Bukti Tidak Mencukupi' dan 'Tidak Berkenaan' dalam Simulasi Mengajar 30 minit serta penilaian semula dalam tempoh dua minggu. Selain itu, Item E7 melibatkan penilaian format dokumen WIM, manakala pemberat markah akhir antara penilai tidak seragam. Berdasarkan AERA et al. (2014) serta Jonsson dan Svingby (2007), ketiadaan deskriptor tingkah laku yang jelas boleh menjejaskan ketekalan penskoran dan ketepatan maklum balas diagnostik sebagaimana dikehendaki dalam Carta Alir Langkah 8.1 Panduan PPI. Dari sudut hujah kesahan Kane (2013), kelemahan ini melemahkan inferens pemarkahan, iaitu pautan pertama yang menghubungkan pemerhatian dengan skor.

Ketiga, dari aspek empirikal dan psikometrik, ketiadaan data skor tempatan yang boleh dirujuk bagi instrumen JPK/P853/D dan JPK/I031/3D-1 (Bidin et al., 2020), serta isu ketidakseragaman penilaian yang diakui dalam Seksyen 1.2 Panduan PPI, menunjukkan keperluan pengujian kesahan dan kebolehpercayaan instrumen secara empirikal. Literatur penilaian prestasi menekankan kepentingan penentukuran penilai bagi meningkatkan ketekalan pemarkahan (Alkharusi et al., 2012; Bell et al., 2012; McCray, 1982; White, 2018). Selain itu, kesesuaian markah lulus minimum 60% perlu diteliti menggunakan kaedah penetapan standard yang sistematik dan boleh dipertahankan (Downing et al., 2006; Tekian & Norcini, 2015). Berbeza daripada kajian terdahulu yang menumpukan pembangunan model pentaksiran TVET pada peringkat makro, kajian ini memfokuskan keperluan rubrik analitik berasaskan Metodologi Pengajaran 4-Langkah yang dilengkapi indikator tingkah laku dan panduan penskoran yang jelas. Justeru, penambahbaikan prosedur pentadbiran sahaja belum memadai tanpa pengujian empirikal terhadap kualiti instrumen dan kesahan keputusan penilaian.

Dapatan ini juga konsisten dengan pengalaman antarabangsa dan tempatan. Gitomer et al. (2021) menunjukkan bahawa penilaian prestasi mengajar guru pelatih yang digunakan secara meluas boleh gagal dari segi reka bentuk pemarkahan dan kebolehpercayaan jika tidak disokong bukti teknikal yang mencukupi. Di Malaysia, Lee et al. (2020) mendapati domain kognitif, inovasi dan refleksi mencatat skor terendah dalam penilaian latihan mengajar, manakala Poli et al. (2024) menggunakan analisis rakaman pengajaran untuk menilai kemahiran menyoal guru pelatih. Kedua-dua kajian ini menunjukkan bahawa kemahiran mengajar boleh dinilai secara terperinci jika indikatornya dinyatakan dengan jelas, sekali gus menyokong cadangan rubrik analitik bagi Simulasi Mengajar TVET-I.

5.2. Spesifikasi DDR, Reka Bentuk Instrumen & Kesinambungan Fasa Kajian Lanjutan

Berdasarkan dapatan analisis keperluan, enam spesifikasi pembangunan instrumen penilaian internship TVET-I dikelompokkan kepada tiga aspek utama, iaitu struktur dan kandungan instrumen, pelaksanaan penilaian, serta pengujian kesahan dan kebolehpercayaan. Rumusan ini selari dengan amalan jaminan kualiti penilaian dalam sistem VET antarabangsa, misalnya panduan *Australian Skills Quality Authority* (ASQA, 2025) yang menekankan alat penilaian yang sah, boleh dipercayai dan disemak secara berkala. Kesahan kandungan dalam Fasa II boleh ditentukan melalui Indeks Kesahan Kandungan (*Content Validity Index*, CVI) oleh panel pakar (Yusoff, 2019), sebagaimana diaplikasikan dalam pengesahan instrumen guru pelatih TVET oleh Abdul Wahab et al. (2025). Dalam kerangka DDR Jenis 1, kajian ini melaporkan fasa analisis keperluan berasaskan dokumen, manakala fasa seterusnya merujuk pembangunan, semakan pakar dan pengujian empirikal instrumen seperti dirumuskan dalam Jadual 6.

Jadual 6. Rumusan Spesifikasi Pembangunan Instrumen Kajian Lanjutan

Kelompok Spesifikasi	Keperluan Reka Bentuk Instrumen	Cadangan Pelaksanaan Kajian Lanjutan (Fasa)
Struktur dan Kandungan Instrumen (Spesifikasi 1 dan 3)	Menetapkan empat konstruk Metodologi Pengajaran 4-Langkah, iaitu Persediaan, Penyampaian, Penggunaan dan Pengesahan. Item pedagogi Bahagian E dipetakan semula, manakala penilaian *WIM* diasingkan. Indikator penilaian disesuaikan mengikut sesi teori (3:2:1) dan amali (1:3:2), dengan mengambil kira tempoh simulasi 30 minit.	Fasa II: Semakan kesahan kandungan oleh 6–10 pakar (Yusoff, 2019) Fasa III: Pengujian struktur konstruk dan kesesuaian indikator teori serta amali.

Kelompok Spesifikasi	Keperluan Reka Bentuk Instrumen	Cadangan Pelaksanaan Kajian Lanjutan (Fasa)
Pelaksanaan dan Penskoran Penilaian (Spesifikasi 2, 4 & 5)	Membangunkan rubrik analitik empat aras dengan deskriptor tingkah laku yang boleh diperhatikan serta kod Bukti Tidak Mencukupi dan Tidak Berkenaan. Instrumen digunakan oleh PP CIAST/CSC dan Penyelia Internship PB untuk tujuan diagnostik dan maklum balas, disokong manual penskoran serta latihan penentukuran penilai.	Fasa II: Pembangunan rubrik, manual penskoran, temu bual kognitif melibatkan 6–8 penilai dan penentukuran penilai. Fasa III: Penilaian kebolehgunaan dan ketekalan pemarkahan antara penilai.
Pengurusan Data dan Pengujian Empirikal (Spesifikasi 6)	Mengasingkan data penilaian pertama daripada penilaian semula, mengurus kod pengecualian tanpa menggantikannya dengan skor sifar serta menganalisis data teori dan amali secara berasingan menggunakan *IBM SPSS Statistics*.	Fasa III: Pengujian kebolehpercayaan dalaman, analisis faktor mengikut kesesuaian data dan penilaian persetujuan antara PP dengan Penyelia PB berdasarkan bukti pemerhatian yang sama.

Secara keseluruhannya, ketiga-tiga kelompok ini menghubungkan dapatan analisis dokumen pada fasa awal dengan keperluan pembangunan dan pengujian instrumen secara sistematik. Pendekatan ini bertujuan menghasilkan instrumen penilaian internship TVET-I yang mempunyai konstruk jelas, panduan penskoran yang seragam serta bukti kesahan dan kebolehpercayaan bagi menyokong penilaian kompetensi pedagogi berasaskan Metodologi Pengajaran 4-Langkah.

5.3. Limitasi Kajian dan Agenda Lanjutan

Kajian ini tertumpu kepada analisis dokumen dasar, instrumen rasmi, rekod penempatan dan sorotan literatur terarah. Oleh itu, lima batasan perlu dinyatakan secara telus bagi memandu kajian lanjutan.

Pertama, himpunan dokumen yang disemak belum merangkumi teks penuh NOSS P853-002-4:2024 (sebagaimana mula dirujuk dalam surat penempatan CIAST bertarikh 29 Julai 2026), pekeliling tarikh kuat kuasa pelupusan VTO Tahap 3 (I-031-3:2014), borang pentadbiran JPK/P853/A, E dan I, serta empat dokumen panduan induk yang dipetik dalam Seksyen 5.0 Panduan PPI (iaitu Panduan Program Pengajaran TVET bertarikh 1 April 2022, Panduan SLaPB COPTPA Edisi II 2024, Panduan Pembangunan Soalan Penilaian Latihan Kemahiran Edisi 2024 dan teks naratif penuh Latihan Internship). Dokumen-dokumen ini perlu diperoleh pada peringkat awal kajian utama bagi mengesahkan sama ada terdapat pindaan kriteria penyampaian selepas edisi 2021.

Kedua, memandangkan tinjauan literatur dalam artikel ini merupakan sorotan terarah awal, masih terdapat kemungkinan wujud literatur empirikal dan teori lain, sama ada di peringkat antarabangsa mahupun tempatan, berkaitan penilaian praktikum guru TVET, rubrik pengajaran mikro, kesetaraan penilaian bersemuka berbanding ulang nilai dan kebolehpercayaan penilai yang belum dirangkumi (Iliasova et al., 2025; Ismail et al., 2017). Justeru, kajian lanjutan akan melaksanakan tinjauan literatur sistematik merentas pangkalan data akademik antarabangsa serta repositori tesis tempatan menggunakan gabungan kata kunci seperti *TVET instructor*, *teaching competence*, *four-step teaching*, *micro-teaching*, *internship*, *practicum*, *performance assessment*, *analytic rubric* dan *rater reliability*.

Ketiga, dapatan ini perlu dilengkapkan dengan triangulasi pengalaman lapangan pihak berkepentingan menerusi temu bual separa berstruktur bersama PP CIAST/CSC, Penyelia PB, penyelaras internship dan pelatih TVET-I bagi menguji kebolehpraktisan pemetaan 43 item dalam Jadual 5, khususnya kedudukan item jambatan seperti B1, B9 dan C4 dalam tempoh sebenar 30 minit.

Keempat, memandangkan Panduan PPI hanya terpakai dalam lingkungan CIAST dan CSC, sebarang peluasan kajian lanjutan kepada penyedia TVET-I di luar rangkaian CIAST memerlukan dokumen peraturan institusi berkenaan serta kebenaran bertulis yang berkaitan.

Kelima, analisis dokumen ini dilaksanakan oleh pengkaji yang terlibat secara langsung dalam ekosistem latihan pengajar (JPK, CIAST dan CSC). Kedudukan orang dalam ini memudahkan akses dan tafsiran dokumen, tetapi berpotensi mempengaruhi tafsiran. Analisis dokumen juga hanya menunjukkan reka bentuk instrumen, bukan tingkah laku sebenar penilai atau taburan skor. Oleh itu, kelemahan yang dikenal pasti dalam kajian ini perlu dianggap sebagai hipotesis pengukuran yang mesti diuji dengan data skor sebenar dalam Fasa III.

6. Kesimpulan

Analisis dokumentasi dalam kajian ini membekalkan bukti berstruktur untuk menjustifikasikan kajian lanjutan bagi membangunkan dan mengesahkan instrumen penilaian pengajaran pelatih internship TVET-I berasaskan Metodologi Pengajaran 4-Langkah. Keperluan kajian lanjutan tidak berpunca daripada ketiadaan borang rasmi TVET-I, sebaliknya daripada gabungan lima kelemahan pengukuran pada borang JPK/P853/D sedia ada, iaitu (1) skala 0–3 tanpa deskriptor tingkah laku; (2) ketiadaan kod pengecualian ‘Bukti Tidak Mencukupi’ dan ‘Tidak Berkenaan’; (3) pertindihan konstruk

WIM menerusi Item E7; (4) item dan wajaran yang seragam merentas sesi teori (3:2:1) dan amali (1:3:2); dan (5) pemberat akhir yang tidak selaras antara penilai. Keperluan ini diperkukuh oleh peningkatan kerumitan kompetensi DKM Tahap 4 berbanding VTO Tahap 3, serta ketiadaan bukti empirikal tempatan mengenai kesahan, struktur dalaman dan ketekalan antara penilai.

Melalui pemetaan kesemua 43 item borang rasmi kepada struktur 4-Langkah dan penelitian enam Unit Kompetensi (CU) NOSS, artikel ini menyediakan asas untuk pembangunan rubrik analitik, pengesahan pakar, temu bual kognitif penilai dan pengujian lapangan menggunakan analisis statistik dalam kajian lanjutan, di samping tinjauan literatur sistematik dan pengesahan dokumen versi 2024. Oleh sebab kajian ini tidak menganalisis data skor sebenar, kelemahan yang dikenal pasti perlu disahkan secara empirikal sebelum sebarang perubahan dibuat terhadap borang rasmi.

Penghargaan

Penulis amat menghargai sokongan, bimbingan dan kerjasama yang diberikan oleh para pengajar dan jurulatih (*coach*) di CIAST dan CSC serta personel JPK, Kementerian Sumber Manusia (KESUMA), dalam menjayakan kajian ini.

Konflik Kepentingan

Penulis mengisytiharkan tiada konflik kepentingan.

Rujukan

Abdul Wahab, Z., Md Osman, S. Z., & Ahmad, N. H. (2025). Pre-service TVET teachers' readiness to apply entrepreneurial elements: Instrument content validity using Content Validity Index (CVI). *Asian Journal of Vocational Education and Humanities, 6*(1), 48–55. https://doi.org/10.53797/ajvah.v6i1.7.2025

Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2012). Educational assessment attitudes, competence, knowledge, and practices: An exploratory study of Muscat teachers in the Sultanate of Oman. *Journal of Education and Learning, 1*(2), 217–232. https://doi.org/10.5539/jel.v1n2p217

Allen, C. R. (1919). *The instructor, the man and the job: A hand book for instructors of industrial and vocational subjects*. J. B. Lippincott.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Antera, S. (2021). Professional competence of vocational teachers: A conceptual review. *Vocations and Learning, 14*(3), 459–479. https://doi.org/10.1007/s12186-021-09271-7

Australian Skills Quality Authority. (2025). *Practice guide: Outcome standards for NVR registered training organisations – Quality Area 1 – Training and assessment: Assessment (Standards 1.3, 1.4, 1.5)* (Version 1.0). https://www.asqa.gov.au/sites/default/files/2025-08/practice-guide-assessment.pdf

Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment, 17*(2–3), 62–87. https://doi.org/10.1080/10627197.2012.715014

Bidin, B., Azman, M. N. A., Tee, T. K., & Khairudin, M. (2020). The quality assessment of teaching delivery based on the National Occupational Skills Standard at MARA High Skills College. *International Journal of Psychosocial Rehabilitation, 24*(9), 274–278.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education, 3*, Article 22. https://doi.org/10.3389/feduc.2018.00022

Che Amran, A., Ananta, G. P., Mat Hanafiah, M. A., Ali, A., & Che Ku Mohd, C. K. N. (2020). Development of the Malaysian Skills Certification for lecturers in tertiary TVET institutions. *Journal of Technical Education and Training, 12*(1), 125–133. https://doi.org/10.30880/jtet.2020.12.01.013

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education, 42*(3), 347–360. https://doi.org/10.1080/02602938.2015.1111294

Downing, S. M., Tekian, A., & Yudkowsky, R. (2006). Procedures for establishing defensible absolute passing scores on performance examinations in health professions education. *Teaching and Learning in Medicine, 18*(1), 50–57. https://doi.org/10.1207/s15328015tlm1801_11

Gitomer, D. H., Martínez, J. F., Battey, D., & Hyland, N. E. (2021). Assessing the assessment: Evidence of reliability and validity in the edTPA. *American Educational Research Journal, 58*(1), 3–31. https://doi.org/10.3102/0002831219890608

Herbart, J. F. (1904). *Outlines of educational doctrine* (A. F. Lange, Trans.; C. De Garmo, Annot.). Macmillan.

Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*(2), 56–64. https://doi.org/10.3102/0013189X12437203

Iliasova, L., Nekrasova, I., Mena, J., & Estrada-Molina, O. (2025). Microteaching on pre-service teachers' education: Literature review. *Frontiers in Education, 10*, Article 1562975. https://doi.org/10.3389/feduc.2025.1562975

Ismail, A., Hassan, R., & Rosli, D. I. (2017). The skill and competency of technical and vocational education and training (TVET) personnel for the development and implementation of a national teacher standard in TVET in Malaysia. *Pertanika Journal of Social Sciences & Humanities, 25*(S), 109–120.

Ismail, K., Mohd Nopiah, Z., Rasul, M. S., & Pang, C. L. (2020). Technical competency among vocational teachers in Malaysian public skills training institutions: Measurement model validation using PLS-SEM. *Journal of Technical Education and Training, 12*(1), 163–174. https://doi.org/10.30880/jtet.2020.12.01.017

Jabatan Pembangunan Kemahiran. (2021). *National Occupational Skills Standard (NOSS) for TVET Instruction Level 4 (P853-002-4:2021)*. Kementerian Sumber Manusia Malaysia.

Jabatan Pembangunan Kemahiran. (2023). *Jurnal Latihan Internship TVET I (P853-002-4:2021 – CIAST – Mac 2023): Borang JPK/P853/B_PP, JPK/P853/B_PPL, JPK/P853/C, JPK/P853/D_PP, JPK/P853/D (PENYELIA) dan JPK/P853/D_PPL*. Pusat Latihan Pengajar dan Kemahiran Lanjutan, Kementerian Sumber Manusia Malaysia.

Jabatan Pembangunan Kemahiran. (2024). *Panduan pembangunan bahan pengajaran dan pembelajaran (BPdP) Edisi 2024* (kemas kini 11 Mac 2025; berkuat kuasa 1 Januari 2025). Kementerian Sumber Manusia Malaysia.

Johnson, E. S., Zheng, Y., Crawford, A. R., & Moylan, L. A. (2020). Examining rater accuracy and consistency with a special education observation protocol. *Studies in Educational Evaluation, 64*, Article 100827. https://doi.org/10.1016/j.stueduc.2019.100827

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130–144. https://doi.org/10.1016/j.edurev.2007.05.002

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Lee, M. F., Lim, S. C. J., & Lai, C. S. (2020). Assessment of teaching practice competency among in-service teacher degree program (PPG) in Universiti Tun Hussein Onn Malaysia. *Journal of Technical Education and Training, 12*(1), 181–188. https://doi.org/10.30880/jtet.2020.12.01.019

Mansor, M., Mohamed, Z., Basir Khan, M. R., & Mohamed Sidek, M. N. A. (2024). Evaluating factors influencing job performance among TVET trainee instructor graduates in Malaysia: A conceptual framework. *SkillsMalaysia Journal, 10*(1), 1–7.

McCray, P. M. (1982). *Vocational evaluation and assessment in school settings* (ED223049). Research and Training Center, Stout Vocational Rehabilitation Institute, University of Wisconsin-Stout. https://files.eric.ed.gov/fulltext/ED223049.pdf

OECD. (2021). *Teachers and leaders in vocational education and training*. OECD Publishing. https://doi.org/10.1787/59d4fbb1-en

Panadero, E., Jonsson, A., Pinedo, L., & Fernández-Castilla, B. (2023). Effects of rubrics on academic performance, self-regulated learning, and self-efficacy: A meta-analytic review. *Educational Psychology Review, 35*(4), Article 113. https://doi.org/10.1007/s10648-023-09823-4

Poli, N. A., Abdul Jobar, N., Mohd Rosli, N., & Janan, D. (2024). Penguasaan tahap kemahiran menyoal dalam pengajaran dan pembelajaran guru pelatih Bahasa Melayu. *Asian Pendidikan, 4*(1), 35–42. https://doi.org/10.53797/aspen.v4i1.4.2024

Pusat Latihan Pengajar dan Kemahiran Lanjutan. (2025). *Panduan tugas dan peranan Pegawai Penilai semasa lawatan internship Program Pengajaran TVET (P853-002-4:2021) bagi Pegawai Penilai CIAST dan Kampus Satelit CIAST* (Edisi 1, kemas kini 15 April 2025). Jabatan Pembangunan Kemahiran, Kementerian Sumber Manusia Malaysia.

Pusat Latihan Pengajar dan Kemahiran Lanjutan. (2026). *Dokumen operasi penempatan dan pemantauan internship TVET-I: Kriteria Pusat Bertauliah Internship dan Penyelia; Senarai Semak Intern; Borang Laporan Kehadiran Latihan Internship (JPK/P853/G); Surat kepada Pengurus Pusat Bertauliah bertarikh 29 Julai 2026; dan Lampiran 1 Jadual Internship Sesi Januari 2026*. Jabatan Pembangunan Kemahiran, Kementerian Sumber Manusia Malaysia.

Richey, R. C., & Klein, J. D. (2007). *Design and development research: Methods, strategies, and issues*. Lawrence Erlbaum Associates.

Saipudin, N. A., & Suhairom, N. (2021). What else matters? Competency of Malaysian educators in technical and vocational education: A scoping review. *Online Journal for TVET Practitioners, 6*(1), 25–31. https://journal.uthm.edu.my/index.php/oj-tp/article/view/7744

Sarkar, M., Gutierrez-Bucheli, L., Yip, S. Y., Lazarus, M., Wright, C., White, P. J., Ilic, D., Hiscox, T. J., & Berry, A. (2024). Pedagogical content knowledge (PCK) in higher education: A systematic scoping review. *Teaching and Teacher Education, 144*, Article 104608. https://doi.org/10.1016/j.tate.2024.104608

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

Smith, E., & Tuck, J. (2023). Do the qualifications of vocational teachers make a difference to their teaching? *Research in Post-Compulsory Education, 28*(1), 1–25. https://doi.org/10.1080/13596748.2023.2166690

Stufflebeam, D. L. (2001). Evaluation models. *New Directions for Evaluation, 2001*(89), 7–98. https://doi.org/10.1002/ev.3

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tekian, A., & Norcini, J. (2015). Overcome the 60% passing score and improve the quality of assessment. *GMS Zeitschrift für Medizinische Ausbildung, 32*(4), Doc43. https://doi.org/10.3205/zma000985

White, M. C. (2018). Rater performance standards for classroom observation instruments. *Educational Researcher, 47*(8), 492–501. https://doi.org/10.3102/0013189X18785623

White, M., & Klette, K. (2024). Signal, error, or bias? Exploring the uses of scores from observation systems. *Educational Assessment, Evaluation and Accountability, 36*(4), 505–528. https://doi.org/10.1007/s11092-024-09427-8

Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education, 58*, 149–162. https://doi.org/10.1016/j.tate.2016.05.010

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal, 11*(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Zulkifli, M. S. (2018). *Model pentaksiran kompetensi holistik dalam pengajaran Pendidikan Latihan Teknikal dan Vokasional (PLTV)* [Tesis doktor falsafah, Universiti Teknologi MARA]. UiTM Institutional Repository. https://ir.uitm.edu.my/id/eprint/82210/

Zulkifli, M. S. (2019). Kaedah pengukuran holistik dalam pembelajaran Pendidikan, Latihan Teknikal dan Vokasional (TVET). *SkillsMalaysia Journal, 5*(1), 16–19.